

SEARCH REQUEST FORM
Scientific and Technical Information Center

Access DB#

57687

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: Sheppard

Searcher Phone #: 308-4499

Searcher Location: _____

Date Searcher Picked Up: _____

Date Completed: _____

Searcher Prep & Review Time: _____

Clencal Prep Time: _____

Online Time: _____

Type of Search

NA Sequence (#) _____

AA Sequence (#) _____

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN _____

Dialog _____

Questel/Orbit _____

Dr. Link _____

Lexis/Nexis _____

Sequence Systems _____

WWW/Internet _____

Other (specify) _____


```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

BE315722
LOCUS BE315722 642 bp mRNA EST 14-JUL-2000
DEFINITION NF025D11LF1090 Developing leaf Medicago truncatula cDNA clone
NF025D11LF 5', mRNA sequence.
ACCESSION BE315722
VERSION BE315722.1 GI:9189499
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabiales; Fabaceae; Papilionoideae; Medicago.
1 (bases 1 to 642)
Flóres-H.R., Inman,J.T., Weller,J.W. and May,G.D.
AUTHORS
TITLES
JOURNAL
COMMENT
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Medicago Genome Initiative accession: MG1:S:16068
Insert Length: 642 Std Error: 0.00
Plate: 025 row: D column: 11
Seq primer: TCACACGAGAAACACGCTATGAC.
FEATURES
source
1..642
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF025D11LF"
/clone_lib="Developing leaf"
/tissue_type="leaf"
/dev_stage="Pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of very
young, developing, mature and senescing leaves."
BASE COUNT 210 a 110 c 109 g 213 t
ORIGIN

Query Match 2.7%; Score 186.8; DB 166; Length 642;
Best Local Similarity 76.2%; Pred. No. 9.5e-22;
Matches 230; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

OY 3502 aggtatgtaagatatagatgaactgtttggagaacacgacgtatcagctgat 3561
|||||
DB 322 AGGTATGTAATTTTCAATACTTGTGGTGACCAACATAGTTTCAAGCAT 361
|||||
OY 3562 gctggacttaatatcatatcacaaaggaagagctcttgaatgtatctc 3621
|||||
DB 382 GCAGATGTAATAGCTCATATTAACAAGGAGGAATATTGATGTCAGTTATCCT 441
|||||
OY 3622 agaagatagtggtgactcttggaagaaagtcgactgcttcttggtagcagcat 3681
|||||
DB 442 AGTAGCTTGTGGTAATCTTGGGAATGCTATGTTGTTGGTGGTGAATGATCAT 501
|||||
OY 3682 agagatatggaagatctcttaacttcttaagtcagcagcagcttgaactatctc 3741
|||||
DB 502 AGAGACTGCTAATATATACCTAATTTTGTGTCATCTAGAGCATCATCTA 561
|||||
OY 3742 cttaaaagatctgagagacactctgtctcttgatctctggcacaacactatc 3801
|||||
DB 562 TTGAAAGAGTTGAGAAACATACCCGCTGTGTTTAAGTCTTGGAAAGAAACCAT 621
|||||
OY 3802 ct 3803
||
DB 622 TT 623
||
RESULT 3

BF324723
LOCUS BF324723 667 bp mRNA EST 21-NOV-2000
DEFINITION su14f03.y1 Gm-cl066 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl066-1781 5', similar to TR:064989 064989 STEROID
22-ALPHA-HYDROXYLASE. [1], mRNA sequence.
ACCESSION BF324723
VERSION BF324723.1 GI:11274370
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabiales; Fabaceae; Papilionoideae; Glycine.
1 (bases 1 to 667)
Shoemaker,R., Kelm,P., Vodkin,L., Erepding,J., Corvelli,V., Khanna
A., Bolla,B., Merrin,M., Hiller,L., Kueba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
R., Ritten,E., Kohn,S., Shin,T., Jackson,T., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
AUTHORS
TITLES
JOURNAL
COMMENT
Public Soybean EST Project
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
High quality sequence stop: 420.
FEATURES
source
1..667
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl066-1781"
/clone_lib="Gm-cl066"
/tissue_type="leaf and shoot tip, salt stressed, 2 week
old seedling"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+, Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from unexpanded leaves and the shoot tips of 2 week old
seedlings from the cultivar Williams. The 2 week old
seedlings were salt stressed in a solution of 500mM NaCl
for 3 days prior to harvesting. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy
Shoemaker."
BASE COUNT 194 a 155 c 141 g 177 t
ORIGIN

Query Match 2.6%; Score 178.6; DB 147; Length 667;
Best Local Similarity 71.4%; Pred. No. 2.3e-20;
Matches 235; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

OY 3502 aggtatgtaagatatagatgaactgtttggagaacacgacgtatcagctgat 3561
|||||
DB 218 AGGTATGTAATTTTCAATACTTGTGGTGACCAACATAGTTTCAAGCAT 277
|||||
OY 3562 gctggacttaatatcatatcacaaaggaagagctcttgaatgtatctc 3621
|||||

Db	278	GCAGACATCAACAGGTCATTCTTACAAAGGAAGGGAATTTGTCAGTGCAGCTATTCCT	337
Qy	3622	agaagatlaagtgatggaattctctgggaaatggtcgaatgctgttcttggtgacatgcac	368
Db	338	AGAAGCATCGGTGGAAATACTAGGAAATATGGTCCATGTTGGTCTTAGTGGTGCATGAT	397
Qy	3682	agagatagaagatcattcgtcttaactcttgaatgaacacgctctgaactatctca	374
Db	398	AGACACTGGGGGTATATATCACTCAACTTCTTAAAGCCACCGCTCAGAACACACTTC	457
Qy	3742	cttaagaatgltgagagacaaactgttgttcttgatctcttggaacaacaaactctatt	380
Db	458	TTGAAAGAGGGTGAAGAGCATATCCCTCTTGATCTTGAACCTCTTGAGCCAAAGATTGCATA	517
Qy	3802	ttctctgtctcaagaacgagcaaaaagt	3830
Db	518	TTCTCAGCCCAAGATGAAGCTAAGAAAGTTT	546

RESULT	4			
LOCUS	BF050501			
DEFINITION	BF050501	533 bp	mRNA	EST 16-OCT-2000
ACCESSION	BF050501	EST353569 tomato developing/immature green fruit Lycopersicon		
VERSION	BF050501	esculentum cDNA clone CLEM16G22.5' sequence, mRNA sequence.		
KEYWORDS	BF050501.1	GI:10804397		
SOURCE	EST.			
ORGANISM	tomato.			
	Lycopersicon esculentum			

REFERENCE	TITLE
AUTHORS	JOURNAL
	COMMENT
1 (Pages 1 to 533) Alcala, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E., Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Romling, C.M., Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley S.D.	Generation of ESTs from tomato fruit tissue, immature green unpublished (2000)
Contact: David Frisch	

100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: df1rsch@clemson.edu.

FEATURES	SOURCE	location/Qualifiers
1.	533	
		/organism="Lycopersicon esculentum"
		/cultivar="T496"
		/db_xref="taxon:4081"
		/clone="CLEM18G22"
		/clone_1bp="tomato developing/Immature green fruit"
		/tissue_type="fruit"
		/dev_stage="Immature green (5-35 days post-anthesis)"
		/lab_host="SOLR"
		/note="Vector: pBluescriptSKmCuaDpT; Site_1: EcoRI; Site_2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and harvested at 7 day intervals through 35 dpa. Equal masses of tissue from each stage were combined (including seeds and locules) prior to mRNA isolation."
BASE COUNT	172 a	
ORIGIN	75 c	175 t

	Query Match	2.5% Best Local Similarity	Score 171.4	DB 143	Length 533
	Matches 220	Conservative 0	Mismatches 81	Indels 0	Gaps 0
Oy	3497	ttcacaggtatgtaagatatagatcgcgaactggtttggagaacacagatcgtatcaag	3556		
Ob	233	ttttcaggatgacggaaaattttcgaagtcacaaattttttggagaacggccaaacaaatgattttc	292		

OY	3557	ctagatgtgacttaataagattcaattttcaaacgaagagagctctttaagt	3616
Db	293	CAGATGAGGGCTAAACAGTACATCTCTGCAGATGAAGGAGATTATTTGAGTAA	352
OY	3617	atccctagaaglatagatgagatctcttggaaatgylcgatgcttgccttglttgtagca	3676
Db	353	ATCCAAAGAAATAGAGGGGATCTGTGTAATAGTCATGTAACTGAAGTTGGACAA	412
OY	3677	tgcctagaaglatagaagaatctgcctctaactcttaagtaagcaagctcttaagata	3736
Db	413	TGCATAGAGTATGAGGATGATTCCCTGTAATTTTTCAGCAATCTAGGCTAAGGAATC	472
OY	3737	tctacttaagaatgcttagagaacatacttgttcttctgattcttgcgacacaact	3796
Db	473	AACCTTAAATGAAGTTGAAGAACATACATTGCTGTCTTGCTGGTGAACACGAT	532
OY	3797		
Db	533	C 533	

[illegible]

REFERENCE
AUTHORS
Shoemaker, R., Keim, P., Vodka, L., Erpelting, J., Coryell, V., Khanna

TITLE	Public Soybean EST Project
JOURNAL	unpublished (1999)
COMMENT	Contact: Shoemaker R/Public Soybean EST Project

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info.genomesystems.com web site: www.genomesystems.com

High quality sequence stop: 388.

Location/Qualifiers
1..522

FEATURES
Source

`/lab host="DH10B"`
`/note="vector: pluscript II XR: Site.1: EcoRI: Site.2:`
`XhoI; This cDNA library was constructed from mRNA isolated`
`from cotyledons of 8-day-old Williams' seedlings which`
`were propagated on paper towels distilled water for`
`3 days (elutriated), then greenhouse grown for 5 days in`
`potting soil. The cotyledons were flash-frozen in liquid`

[illegible][illegible]

LOCUS	AM398669	615 bp	mRNA	EST	07-FEB-2000
DEFINITION	EST309169 L. pennellii trichome, Cornell University Lycopersicon pennellii cDNA clone cLPT4K22 5', mRNA sequence.				
ACCESSION	AM398669				
VERSION	AM398669.1 GI:6917139				
KEYWORDS	EST.				
SOURCE	Lycopersicon pennellii.				
ORGANISM	Lycopersicon pennellii. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.				
REFERENCE	1 (bases 1 to 615) Alcala,J., Vrebalov,J., White,R., Matera,A.L., Lakey,J., Holt,I.E., Liang,F., Hansen,T.S., Upson,J., Rooning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Niernan,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and Giovannoni,J. Generation of ESTs from wild tomato (Lycopersicon pennellii) trichomes				
TITLE	unpublished (1999)				
JOURNAL	Contact: David Frisch Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfirsch@CLEMSON.EDU 5 prime sequence.				
FEATURES	Location/Qualifiers 1..615 /organism="Lycopersicon pennellii" /db_xref="taxon:28526" /clone="cLPT4K22" /clone_1lb="L. pennellii trichome, Cornell University" /tissue-type="trichome" /dev_stage="mixed stages" /lab_host="SOLR" /note="vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Leaves of various stages were shaken in liquid nitrogen, shearing off trichomes. This procedure yielded a mixture of cells highly enriched for trichomes, with minor contamination by other types of leaf cells."				
BASE COUNT	180 a 111 c 100 g 224 t				
ORIGIN					
Query Match	1.5%; Score 105; DB 115; Length 615; Best Local Similarity 57.0%; Pred. No. 5.1e-08; Matches 192; Conservative 0; Mismatches 145; Indels 0; Gaps 0;				
QY	3496 ttcttaaggtatggaagatatatgatgcagcactggttggaagaacagatglatca 3555				
Db	265 tcttttaggtatggagaaagtgttcaagtcacacttattttttctcccaacagtggtca 324				
QY	3556 gctgtagctgagctcaatagatcatatcacaaaaggaaggctctcttgaaatgagt 3615				
Db	325 ttgtgaccaaagaccttaattacttcatattttacaaatgaagataagttatttcagctagt 384				
QY	3616 tatctctagaagratatggttggagattctctgggaaatgctcgatgctctctgttgagc 3675				
Db	385 tatccaaagcccaattcatatggtttactttggcaaaagttttcatttgcttgctgttgctac 444				
QY	3676 atgcagtagatatagagaagatctcgcgtctaactcttaagtaacgacgctctagaact 3735				
Db	445 acacgtataaaagccttagacatgttttcattatnaccatnaccagacacatttaagtcttaacct 504				
QY	3736 attactactaaagatgcttgagagacatacttgtttctcttgatctcttgagcaacaaac 3795				
Db	505 gagtttatttaagatgttgaaacatttagacacttgagattcttcattcatatggaagatmaa 564				
QY	3796 tctatttctctgcctcaagaagagccaaaaggttc 3832				

RESULT	13
LOCUS	BF052051
DEFINITION	BF052051 191 bp mRNA EST 16-OCT-2000
ACCESSION	EST437298 tomato developing/immature green fruit Lycopersicon
VERSION	esculentum cDNA clone cLEM25M4 5' sequence, mRNA sequence.
KEYWORDS	BF052051
SOURCE	BF052051.1 GI:10805947
ORGANISM	EST.
SOURCE	tomato.
REFERENCE	Lycopersicon esculentum
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 191) Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Romling,C.M., Niernann,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tankeley ,S.D. Generation of ESTs from tomato fruit tissue, immature green unpublished (2000) Contact: David Frisch Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrisch@CLEMSON.EDU. Location/Dualifiers 1..191 /organism="Lycopersicon esculentum" /cultivar="TRA96" /db_xref="taxon:4081" /clone="cLEM25M4" /clone_1lb="tomato developing/immature green fruit" /tissue_type="fruit" /dev_stage="Immature green (5-35 days post-anthesis)" /lab_host="SOLR" /note="Vector: pBluescriptSKmCuAdapt; Site:1: EcoRI; Site:2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and harvested at 7 day intervals through 35 dpa. Equal masses of tissue from each stage were combined (including seeds and locules) prior to mRNA isolation."
BASE COUNT	65 a 19 c 46 g 61 t
ORIGIN	
Query Match	1.5%: Score 103.6; DB 143; Length 191;
Best Local Similarity	71.6%; Pred. No. 8.9e-08; Gaps 0;
Matches 136; Conservative	0; Mismatches 54; Indels 0;
OY	3574 agatcatattacaaagaaggaaggctttgagttagtactcctaagaataagtt 3633
Db	2 AGATCATATTGCAAGAAAGAAGGAGATTATTGTAGTGAATTATTCCAAAGATTAGCT 61
OY	3634 gggatctctgggaatatgttcgatcgctgttcttggttgtagacatgcataagatatgga 3693
Db	62 GGGATACTTTGGTAATVGGCTATGTTAATTCAAGTTGGACAATAATGCATATGATGAGG 121
OY	3694 agtatctcgcttaaccttctaagtaacgcacgctcttagaacattactactaaagatgtt 3753
Db	122 ATGATTTCTCGAATTTTITTAGCAATGCTATGCTAAGATTCACACTTTTAAGTGCAATT 181
OY	3754 gagagacata 3763
Db	182 GAAGAAGCAT 191
RESULT	14
LOCUS	FM615836

Mon Oct 1 11:34:40 2001

us-09-502-426-1.rst

Page 12

Search completed: September 29, 2001, 11:18:30
Job time: 16099 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2001, 06:50:21 : Search time 139.83 Seconds
(without alignments)
9325.431 Million cell updates/sec

Title: US-09-502-426-1

Perfect score: 6888

Sequence: 1 atgttggtatttatgttg.....cctagtcagccaagcct 6888

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCrUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92.8	1.3	4937	2	US-08-622-166A-3
2	87.6	1.3	1608	2	US-08-622-166A-1
3	77.4	1.1	19124	2	US-08-487-826B-13
4	68.6	1.0	837	4	US-08-998-416-288
5	67.4	1.0	19124	2	US-08-487-826B-13
6	65.8	1.0	615	4	US-08-998-416-186
7	65.4	0.9	636	4	US-08-998-416-1137
8	63.6	0.9	8920	2	US-08-446-855A-1
9	63.6	0.9	8920	2	US-09-150-741-1
10	62.6	0.9	4818	3	US-08-817-926-27
11	62.6	0.9	7218	1	US-08-232-463-14
12	61.8	0.9	5852	1	US-07-867-106-2
13	60.8	0.9	4818	3	US-08-817-926-27
14	60.6	0.9	837	4	US-08-998-416-288
15	60.4	0.9	8920	2	US-08-446-855A-1
16	60.4	0.9	8920	2	US-09-150-741-1
17	60.2	0.9	665	2	US-08-883-795A-36
18	59.2	0.9	51952	3	US-08-947-823-1
19	59	0.9	827	4	US-08-998-416-535
20	58.4	0.8	665	2	US-08-883-795A-36
21	57.8	0.8	1511	1	US-07-991-867B-8
22	57.8	0.8	1511	1	US-08-107-755A-8
23	57.8	0.8	1511	1	US-08-544-332-8
24	57.4	0.8	3095	6	5231168-1
25	56.4	0.8	2058	2	US-08-749-391-1
26	56.4	0.8	2058	3	US-09-390-200-1
27	56.4	0.8	4673	1	US-07-638-431-1

28	56.4	0.8	4673	5	PCT-US92-00018-1	Sequence 1, Appli
29	56.2	0.8	2960	3	US-08-913-842-3	Sequence 3, Appli
30	55.6	0.8	701	4	US-08-998-416-701	Sequence 701, App
31	55.6	0.8	5852	1	US-07-867-106-2	Sequence 2, Appli
32	55.2	0.8	1406	3	US-08-913-842-6	Sequence 6, Appli
33	55.2	0.8	1511	1	US-07-991-867B-8	Sequence 8, Appli
34	55.2	0.8	1511	1	US-08-107-755A-8	Sequence 8, Appli
35	55.2	0.8	1511	2	US-08-544-332-8	Sequence 8, Appli
36	55	0.8	1850	3	US-08-617-860B-32	Sequence 32, Appli
37	55	0.8	4098	2	US-08-605-106-4	Sequence 1, Appli
38	54.8	0.8	6768	1	US-08-107-755A-1	Sequence 1, Appli
39	54.8	0.8	8457	1	US-07-991-867B-1	Sequence 1, Appli
40	54.8	0.8	8457	2	US-08-544-332-1	Sequence 1, Appli
41	54.6	0.8	9048	3	US-08-973-273-4	Sequence 4, Appli
42	54.2	0.8	663	4	US-08-998-416-191	Sequence 191, App
43	53.8	0.8	6243	2	US-09-056-075-1	Sequence 1, Appli
44	53.6	0.8	2430	4	US-08-845-258-3	Sequence 3, Appli
45	53.6	0.8	2430	4	US-08-990-571-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-622-166A-3
Sequence 3, Application US/08622166A
Patent No. 5952545
GENERAL INFORMATION:
APPLICANT: KONCZ, CSABA
APPLICANT: MATHUR, JADEEP
APPLICANT: SZEKERES, MIKLOS
APPLICANT: ALTMANN, THOMAS
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/622,166A
FILING DATE: 27-MAR-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 0147-0153P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4937 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
STRAIN: cv. Columbia
IMMEDIATE SOURCE:
LIBRARY: lambda gt10

CLONE: C204
FEATURE:
NAME/KEY: CDS
LOCATION: join(968..1483, 1680..1829, 1917..2165, 3903
LOCATION: ..3989, 4084..4162, 4248..4354, 4446..4576, 4674
LOCATION: ..4773)
US-08-622-166A-3

Query Match 1.3%; Score 92.8; DB 2; Length 4937;
Best Local Similarity 53.1%; Pred. No. 2.1e-09;
Matches 225; Conservative 0; Mismatches 192; Indels 7; Gaps 1;

QY 3501 caggtatgtagatataatagatgaactgttttgagaagaacagatcgtatcagctga 3560
DB 1162 CCGGTACGGTTCGGTTTCATGACGATCTTTTGGTGAACCGACATTTTCACCTGA 1221
QY 3561 tgcgtacttaataagattcaatatacaaaaggaagagctcttgatgtatctc 3620
DB 1222 CCGGGAAGACGCGGTTTGTCTTTCAGAACGAGGAGCTTTTGAAGTGTCTTATCC 1281
QY 3621 tagaagtagagtgaggtctcttgagaatggtcagatgctgtctgtgtgtagatga 3680
DB 1282 TCGTTCATTTTGAACCTTTTGGGGAACACCTCTGCTTCTTAAAGGTTCTTTGCA 1341
QY 3681 tagagataagagaagatctcgtcttaactcttaagtagcaagcgtcttagaacttct 3740
DB 1342 TAAACGTATGCACTCTCCTACCATGAGCTTTGCTAATTCCTCAATCAATCAAGCATCT 1401
QY 3741 acttaagatgtagagagacatactctgttctgtctgtctgtcgaacaacatctat 3800
DB 1402 CATGCTGATATTGACCGGTTACTCGGTTTAACTTGATCTTGG-----TCTTCTC 1454
QY 3801 ttctctgtcgaagaagagccaaaaggttttaatttattatcttatttgcataatt 3860
DB 1455 GTGTTCTCCATGAGAAAGCAAAAGGTAAACAAAATTCGTATCAAAAAC 1514
QY 3861 ttctgttataagactttagagttcttaactttttttttaaattgaagcttaagct 3920
DB 1515 ATATATATATATATATATATGCGCTTCTCATCTATGTTTTTAAATAAATAAATAAT 1574
QY 3921 ttaa 3924
DB 1575 AAAA 1578

RESULT 2
US-08-622-166A-1
Sequence 1, Application US/08622166A
Patent No. 5952545

GENERAL INFORMATION:
APPLICANT: KONCZ, CSABA
APPLICANT: MATHUR, JAIDEEP
APPLICANT: SEKERES, MIKLOS
APPLICANT: ALTMANN, THOMAS
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME
NUMBER OF SEQUENCES: 4
P450-TYPE PROTEINS INVOLVED IN THE BRASSINOSTEROID SYNTHESIS
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/622,166A
FILING DATE: 27-MAR-1996

CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 0147-0153P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1608 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: A. thaliana

IMMEDIATE SOURCE:
LIBRARY: lambda gt10
CLONE: C204
FEATURE:
NAME/KEY: CDS
LOCATION: 48..1466
US-08-622-166A-1

Query Match 1.3%; Score 87.6; DB 2; Length 1608;
Best Local Similarity 56.6%; Pred. No. 1.6e-08;
Matches 162; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 3501 caggtatgtagatataatagatgaactgttttgagaagaacagatcgtatcagctga 3560
DB 242 CCGGTACGGTTCGGTTTCATGACGATCTTTTGGTGAACCGACATTTTCACACTGA 301
QY 3561 tgcgtacttaataagattcaatatacaaaaggaagagctcttgatgtatctc 3620
DB 302 CCGGGAAGACGCGGTTTGTCTTTCAGAACGAGGAGCTTTTGAAGTGTCTTATCC 361
QY 3621 tagaagtagagtgaggtctcttgagaatggtcagatgctgttctgtgtgtagatga 3680
DB 362 TCGTTCATTTTGAACCTTTTGGGGAACACCTCTGCTTCTTAAAGGTTCTTTGCA 421
QY 3681 tagaagtagagaagatctcgtcttaactcttaagtagcaagcgtcttagaacttct 3740
DB 422 TAAACGTATGCACTCTCCTACCATGAGCTTTGCTAATTCCTCAATCAATCAAGCATCT 481
QY 3741 acttaagatgtagagagacatactctgttctgtctgtatcttg 3786
DB 482 CATGCTGATATTGACCGGTTAGTCCGTTTAACTTGATCTTGG 527

RESULT 3
US-08-487-826B-13
Sequence 13, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chiltons, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th floor
CITY: Newport Beach
STATE: California

COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 1.1%; Score 77.4; DB 2; Length 19124;
Best Local Similarity 42.6%; Pred. No. 2.9e-06;
Matches 590; Conservative 0; Mismatches 781; Indels 14; Gaps 3;

QY 1013 ttaaaacgattttacacgagcactagltgagattctagctagcaataagcaataagcacc 1072
DB 229 TTATATATCAATATATATATATATATATATATATATATATATATATATATATATAT 288
QY 1073 gttcaagctatttacaagaagttacaactgtaataagcttgaatacccttgaataatt 1132
DB 289 TTCTTGTTATTTTATATATATATATATATATATATATATATATATATATATATATAT 348
QY 1133 tggaaatccggttgctgtaataatagatttagtgtaacaataatgtaatacaatt 1192
DB 349 AATTCTTATATCTTTTATGCAACAAAAACATTAAGTAATTCATACATACACAAAA 408
QY 1193 agtgcacacataacataatccttcaagaacaaacaaacttaagagaagttacatat 1252
DB 409 AAAAAAATAAAAAAATAAAAAATTTATATATATATATATATATATATATATATATAT 468
QY 1253 ccataataggtatgctatccttccctacgtatgctatagactaagaataagttat 1312
DB 469 TCACCTATATATATATATATATATATATATATATATATATATATATATATATATAT 528
QY 1313 gtgagtcgataaataagaaatcacacgctggttaataataatgaggaacgtaatacagat 1372
DB 529 TTTAATCATGGAAGAGTTAAGAAATCATTTTATTTTATTTTATTTGATATTCATCAACT 588
QY 1373 cactgcacaaatcattctgtgtgctcaacaataaacaacaaacagaacaaagaanaa 1432
DB 589 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 648
QY 1433 cgattttcttgattcattcattgataatgaatgataagcttcttggttacaagttt 1492
DB 649 TCTTTTATC-----ACATGGAATATTTAAATAATTTTCGATTTATTCGATATATTT 702
QY 1493 cgaagtcctctacaagcggttaacacatcgcaacttaaatgcttcttctaagcacc 1552
DB 703 ATGTGTTATATATATATATATATATATATATATATATATATATATATATATATATAT 762
QY 1553 tttaacataattatgttagttggaatttaataagagcgaactgttaacattacatat 1612
DB 763 AATTAATATTAATAATCTGATTTATTAATTCATTTATATATATATATATATATATATAT 822

QY 1613 tatattagatactagtagtatttccaaataacataacttgatgtttaactaatc 1672
DB 823 TTATTTTCAAAAATCAATTAATTAAGATTTCTTAAATAATTTCTCAATTTATTCATTTTATA 882
QY 1673 ttgttcttccacggtataaataatcatcgcagggtaaaaaagtttgcttattt 1732
DB 883 GCATATACATGCACATTTTAAATTTTATATAAAAAATTTTATTTTATATATATATATATACA 942
QY 1733 cgcagatgataaggaataaaccttaatgaatttcttgaataatgtaaccccttact 1792
DB 943 ATTTTCATACATTAATTTTTCACACACATTTAAGTTGTCTAATATGTAACCATTAAT 1002
QY 1793 catagataattacccgtatgttctgtgcataatgacagccttacaacgtgatagt 1852
DB 1003 AATTAATATCTTATATATATATATATATATATATATATATATATATATATATATATAT 1062
QY 1853 caattttcttgcataataataatgaataatcaatgctactatacagaagaacagc 1912
DB 1063 CTGATTTATCATAGATATATATATATATATATATATATATATATATATATATATATAT 1122
QY 1913 tgaattacacatttaatttaagaacaaatttggaaaatgtaataattcttaacaat 1972
DB 1123 TTAATAAATATATATATATATATATATATATATATATATATATATATATATATATAT 1182
QY 1973 attataaataatgtagtactaagatgttccctatgttcttaaaatattttttata 2032
DB 1183 TAAGATATAATCTTATATATATATATATATATATATATATATATATATATATATATAT 1237
QY 2033 tttagttataacatacattatgacacataatagttgggtgaattcaaatccataat 2092
DB 1238 TTATATATATAGTATATATATATATATATATATATATATATATATATATATATATATAT 1294
QY 2093 atttttgaacatcaaatlaataatattgtaataacaaatgcatagaagaagttccaa 2152
DB 1295 TTTGTATTAACATATCAATCAAGACTTAAGAAACTATATCAATCTGATCTATATATATAT 1354
QY 2153 aaaaaatttgttacaagaacttccaaatttttttttttttttttttttttttttttttttt 2212
DB 1355 ATTAATATCTTTTATATATATATATATATATATATATATATATATATATATATATATAT 1414
QY 2213 atagaanaactatttctgttgcgaagtagtaataacataacataacataacataacata 2272
DB 1415 TAATATATATTTTTCATATATATATATATATATATATATATATATATATATATATATAT 1474
QY 2273 attataaagccttaacgctcaagtaatgtaataatgtaggtgtaataatgtaataatgcat 2332
DB 1475 AAAACCTTCAAAACATTTTTCATTAATAATTAATATATATATATATATATATATATAT 1534
QY 2333 ggtgcagatcagaattgggacaacatggaacggaataataataataataataataata 2392
DB 1535 AGAGAGAAACGTAGAACATACCAAAAAAATTAAGACAAAAAAGAAATATATACAAAAATTA 1594
QY 2393 ataaa 2397
DB 1595 TAAAA 1599

RESULT 4
US-08-998-416-288
; Sequence 288, Application US/08998416
; Patent No. 6239264
GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Redischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYIPII
; NUMBER OF SEQUENCES: 1152


```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/998,416
: FILING DATE: 24-DEC-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: CH 0016/97
: FILING DATE: 31-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8587
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 186:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 615 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: PAC1074RP
:
: US-08-998-416-186

```

```

Query Match          1.0%; Score 65.8; DB 4; Length 615;
Best Local Similarity 48.3%; Pred. No. 0.00021;
Matches 216; Conservative 0; Mismatches 227; Indels 4; Gaps 1;

```

```

Db 1750 aaacctaatgcttaatttttgaataatgtaaccttttctaatagataattaccgt 1809
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
56 ATATTAAAGTTAAATTAATTTAAACGATTTATCATTAATTAATTAATTAATTAATTTG 115

Qy 1810 atgttttgcacatgacgcctctacacactgtagatgacatttttctgcaaat 1869
      || || || || || || || || || || || || || || || || || || || ||
Db 116 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 175

Qy 1870 attaatgaagaatccaatgctactacatcaatgaagaacagctgagatataattaa 1929
      || || || || || || || || || || || || || || || || || || || ||
Db 176 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 235

Qy 1930 tttaagacacaaatttttgaataatgtaataatttctaacataattttaaataatg 1989
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 ATTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 295

Qy 1990 cctaatagttctcctaagcttaataataattttt---cttaataattagttataa 2045
      || || || || || || || || || || || || || || || || || || || ||
Db 296 ATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 355

Qy 2046 acattatgaacacaaatagctgtagaattccaataatccatcaataatttttgaatc 2105
      || || || || || || || || || || || || || || || || || || || ||
Db 356 AAGATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 415

Qy 2106 tacaataattatattagtcacataacatgcatagagaattccaaaaaaatttgg 2165
      || || || || || || || || || || || || || || || || || || || ||
Db 416 AATATTATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 475

Qy 2166 aacgaacactccaattttttttt 2192
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 AACCTTAATTTCTTAATTAATTAATTTT 502

```

```

RESULT 7
US-08-998-416-1137
: Sequence 1137, Application US/08998416
: Patent No. 6239264
: GENERAL INFORMATION:
: APPLICANT: Philippsen, Peter
: APPLICANT: Pohlmann, Rainer
: APPLICANT: Steiner, Sabine
: APPLICANT: Mohr, Christine
: APPLICANT: Wendland, Jurgen

```

```

: APPLICANT: Knechtle, Philipp
: APPLICANT: Reibischung, Corinne
: TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYPII
: TITLE OF INVENTION: AND USES THEREOF
: NUMBER OF SEQUENCES: 1152
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 6239264artis Corporation
: STREET: 3054 Cornwalis Road
: CITY: Research Triangle Park
: STATE: No. 6239264th Carolina
: COUNTRY: USA
: ZIP: 27709
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/998,416
: FILING DATE: 24-DEC-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: CH 0016/97
: FILING DATE: 31-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8587
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 1137:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 636 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: PAC1692RP
:
: US-08-998-416-1137

```

```

Query Match          0.9%; Score 65.4; DB 4; Length 636;
Best Local Similarity 47.1%; Pred. No. 0.00025;
Matches 269; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

```

```

Qy 1841 aacgtgatagcatttttctgcaaatatataatgaagaattcgaatgctactacaat 1900
      || || || || || || || || || || || || || || || || || || || ||
Db 29 AAGATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 88

Qy 1901 agaagaacagctgtagatcatttcaatttaagaacaaattttgaaatgtaata 1960
      || || || || || || || || || || || || || || || || || || || ||
Db 89 ATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 148

Qy 1961 attcctaacaatattatataa-aataatgctcctaataatgtatccatgcttctaaat 2019
      || || || || || || || || || || || || || || || || || || || ||
Db 149 TTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 208

Qy 2020 atttttttattatagttatataatcatatgacaacaaatagctgtagaattcaaa 2079
      || || || || || || || || || || || || || || || || || || || ||
Db 209 ACTATTATGCTATGCTCAAAATTTAAATTTAGTTATTAATTAATTAATTAATTAATTA 268

Qy 2080 tatctcatataatttttgaatctacaatattatataatattagtcataacaatgca 2139
      || || || || || || || || || || || || || || || || || || || ||
Db 269 TTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 328

Qy 2140 tagaagttccaaaaaaattttgtaacgaacactccaaatttttttttcaatgaa 2199
      || || || || || || || || || || || || || || || || || || || ||
Db 329 ATTAAA-----ATATATATTATTATTATTAAAGATTTAATTAATTAATTAATTAAT 383

Qy 2200 caagaataacagatagaataactatttggtagtgagtgaggaatataataatgaag 2259
      || || || || || || || || || || || || || || || || || || || ||

```


Db 384 TTATTTATTTTATTAATAATTCATTTTATAAATAATATGTTGATTAATATATTTAT 443

Qy 2260 Caaatttcaaaaaataataaagcctatacgcgcctcaaaagtlytbatctagtggytla 2319

Db 444 CTTTTTTAAAGAAATTTATTTAAAAATTAATTTAACTTTAAATTCCTATTTATTTATTTT 503

Qy 2320 attataatgcaggygcggttcaagaattgggcaacaatlgaaaacggaaatlaaataatt 2379

Db 504 ATATTATTTAATAAATATATTCATTTTATTTATTTATTTAATAAATAATTTATT 563

Qy 2380 aacttaaaataaataaaattcgaglaaat 2410

Db 564 TTAATTAATATTTTATTCATTTTAAATTAAT 594

RESULT 8
US-08-446-855A-1
; Sequence 1, Application US/08446855A
; Patent No. 5840573

1 APPLICANT: Flores, Maria V
 2 APPLICANT: O'Sullivan, William J
 3 TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
 4 TITLE OF INVENTION: phosphate synthetase II
 5 NUMBER OF SEQUENCES: 2
 6 CORRESPONDENCE ADDRESS:
 7 ADDRESSEE: Nixon & Vanderhye PC
 8 STREET: 1100 No. 5849573th Glebe Road, 8th Floor
 9 CITY: Arlington
 10 STATE: Virginia
 11 COUNTRY: USA
 12 ZIP: 22201-4714
 13 COMPUTER READABLE FORM:
 14 MEDIUM TYPE: Floppy disk
 15 COMPUTER: IBM PC compatible
 16 OPERATING SYSTEM: PC-DOS/MS-DOS
 17 SOFTWARE: PatentIn Release #1.24
 18 CURRENT APPLICATION DATA:
 19 APPLICATION NUMBER: US/08/446,855A
 20 FILING DATE: 06-Jul-1995
 21 CLASSIFICATION: 435
 22 ATTORNEY/AGENT INFORMATION:
 23 NAME: Mitchard, Leonard C
 24 REGISTRATION NUMBER: 29.009
 25 REFERENCE/DOCKET NUMBER: 47-80
 26 TELECOMMUNICATION INFORMATION:
 27 TELEPHONE: 703-816-4000
 28 TELEFAX: 703-816-4100
 29 INFORMATION FOR SEQ ID NO: 1:
 30 SEQUENCE CHARACTERISTICS:
 31 LENGTH: 8920 base pairs
 32 TYPE: nucleic acid
 33 STRANDEDNESS: single
 34 TOPOLOGY: linear
 35 MOLECULE TYPE: genomic
 36 US-08-446-855A-1

Query Match	0.9%;	Score 63.6;	DB 2;	Length 8920;
Best Local Similarity	46.8%;	Pred. No. 0.0011;		
Matches 278; Conservative	0;	Mismatches 304;	Indels 12;	Gaps 2

[illegible]

D	b	272	TATATTTTAACAACAAACGTTTTAGCTGTATTTATATATATATATATATATATA	331
O	y	2053	gaaccaataaatgttggtaattcaaaatctccattaaatlcttttgaaactcaaat	2112
D	b	332	TATATATATATATATGTTTGTCGTCATTTGTTTAAAATTAAGTGAATAATAAACT	391
O	y	2113	tattaatatlttaglcaaataacaaigtacaagaaagtcccaaaaaaatlgtlaacaga	2172
D	b	392	TATTAATATATTTCCCAATTAATATGAATCAATATTATATTTTGATCGTGCACATT	451
O	y	2173	aactccaatlctlltlctlatgga-----acaagaataacagatgsaaaac	2221
D	b	452	ATATAGTTTACACTCCTCTATATATAAACACATCCATATATTTATCACAAATATATAATAC	511
O	y	2222	tatttgttbggaatvgaaagtgtcatatcaccttaagcgaatlctaanaatlatala	2281
D	b	512	TCCCCAATATGGTGTCCTATATTTTATATATATTTATTTATTTATTTATTTATTCATT	571
O	y	2282	gccttatcgsgclcaaaagtgtgatctatccttaggtgtaattataatgaatggtgcgatl	2341
D	b	572	ATTTATTTTTTTTCTTAGGTTATPAATAATAGTAATCTACTAATTTAAAAAATTTTTAAAAA	631
O	y	2342	cagaattcggaacaacaaatgaaaaocgaattaaaatactaacitaaaatnaataaaaatl	2401
D	b	632	AAAAAAAAAAAAAAAAAAAAAAAAAAATTTACATATGAAAAAAGAACCCTGTATATGTPAA	691
O	y	2402	tgaatgaatgttgtcttgactatgaggggcataaanaagaacatgycanaag	2455
D	b	692	TTTTTAAATATTTTAAACATTAATATATATATGCTTAAAAAATTTTTAAAAAATG	745

RESULT 9
US-09-150-741-1
; Sequence 1, Application US/09150741

```

1  APPLICANT: Stewart et al.
2  TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
3  Patent No. 6183996
4  TITLE OF INVENTION: Synthetase II
5  FILE REFERENCE:
6  CURRENT APPLICATION NUMBER: US/09/150,741
7  CURRENT FILING DATE: 1998-09-10
8  EARLIER APPLICATION NUMBER: PL6380
9  EARLIER FILING DATE: 1992-12-16
10 EARLIER APPLICATION NUMBER: A093/00617
11 EARLIER FILING DATE: 1993-12-02
12 EARLIER APPLICATION NUMBER: 08/446,855
13 EARLIER FILING DATE: 1995-07-06
14 NUMBER OF SEQ ID NOS: 15
15 SOFTWARE: PatentIn Ver. 2.0
16 SEQ ID NO 1
17 LENGTH: 8920
18 TYPE: DNA
19 ORGANISM: Plasmodium falciparum
20 US-09-150-741-1

```

Query Match	0.9%	Score 63.6	DB 4	Length 8920
Best Local Similarity	46.8%	Pred. No. 0.0011		
Matches 278	Conservative 0	Mismatches 304	Indels 12	Gaps 2

[illegible]

QY 2053 gaaccataatagtgttgatgcataaataccocattataatttttgaataatcccaaat 2112
 Db 332 tatataataataatgtttgtgtgttcattgttttaataatctctgaaataataaact 391
 QY 2113 tathaataatttaglcaataacaaatgcataagaaagttcccaaaaaaattttgtaacgaa 2172
 Db 392 tattaataataattccaatttaataatagatacaattataattttgtgtgtaacaccta 451
 QY 2173 acttcccaatttttttttttattgga-----acaagaataacagatagaanaac 2221
 Db 452 atatagatttaccattctcttaataaacaacccataatattatatacaatatataaac 511
 QY 2222 taattgttggaatggaagatgaatgaataacataagaanaattttaaanaatatataa 2281
 Db 512 tcccacaatacttggttctcctaataattttatatatatatttatttataattatcatt 571
 QY 2282 gacctatacgctcccaagatagtattatctagtgtgaattatataatgatctgtcgatt 2341
 Db 572 atttatttttttctctagtattataaaaatgtaattcttacttaatttaaaaaaaaaa 631
 QY 2342 cagaatctgggacaacaatgaaaaacggaattaaataatttaactttaataataaanaatt 2401
 Db 632 aaaaaaaaaaaaaaagaaaaaaaataatttacaatgaaaaatgaaactgtcatagttaa 691
 QY 2402 tgaagtaaatgtgtttctgactcttgaggggcaaaaaaagaacatgtccaaag 2465
 Db 692 ttataataatttttaaacataataataatgtatcaaaaaaagaaaaaagt 745

US-RESULT 10
US-08-817-926-27
Sequence 27, Application US/08817926
Patent No. 6001590
GENERAL INFORMATION:
APPLICANT: Komeda, Toshihiro
APPLICANT: Suda, Hisako
APPLICANT: Tamai, Yukio
APPLICANT: Imawatsu, Akihito
APPLICANT: Kato, No. 6001590uo
APPLICANT: Sakai, Yasuyoshi
TITLE OF INVENTION: PROMOTER/TERMINATOR FOR CANDIDA BODINI
TITLE OF INVENTION: FORMATE DEHYDROGENASE GENE
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,926
FILING DATE: 09-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/02597
FILING DATE: 12-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 234133/1995
FILING DATE: 12-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 42536/1996
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 081356/0112

```

? TELECOMMUNICATION INFORMATION
? TELEPHONE: (202)672-5300
? TELEFAX: (202)672-5399
? TELFX: 904136
? INFORMATION FOR SEQ ID NO: 27:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4818 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? ORIGINAL SOURCE:
? ORGANISM: Candida boldinii
US-08-817-926-27

```

Query Match	0.98;	Score 62.6;	DB 3;	Length 4818;
Best Local Similarity	44.48;	Pred. No. 0.0015;		
Matches 338;	Conservative 0;	Mismatches 419;	Indels 4;	Gaps 2

QY	1525	actcttaaatctgctctcttaagaacatcttaacgactcttaacattctgttgatgctgaattcaat	1584
Db	4020	AACTCTTAAATGATATCTTTTATTGATATTTTTCATATATTTCTTGATTAATAATTTGTG	4079
QY	1585	aagaagcgaactgyltaacaatlacaaatlatalatagatactagatgylgatlattcccaa	1644
Db	4080	TTGAAATTTCAATGATATATTTATTTCTTAGATTTTAAAAAAATATTTTCCCTCT -TTT	4138
QY	1645	tacacttgagtgltttaactaactctgttcttctctacggatataataatctaat	1704
Db	4139	TATTTTCCTTTGGTTTTTCCAAATTTTAAATGATAGANTGAAAGAAATTAATAATATATTT	4196
QY	1705	cgaagtaaaaaagctcttgctcattcttcgcgatgaagaalaaacctaaactgacct	1764
Db	4199	ATTATATATTTATTTTACTGCAACCTTTTAAAAAGAGATGATTTAACTTAACACCTCT	4256
QY	1765	aattcttgaaaaatglaacccttact--catagatbaatlaccglatglttgctg	1821
Db	4259	CTGAGTTATATATGATATAAACAGTGTATTAATATACCTTTTTCCTCAAAATTTTTCGA	4318
QY	1822	ccataatgaagcctctacacgcgtgatgtatctatlltcttcgcaaatataatlag	1881
Db	4319	TTATATCATTTTGGTATTTGGTTAATTTATGTCGAATTAATAAAAAAGAAAAATTTCTGT	4378
QY	1882	atccaatgcctacatcaatagaagaacagcgagatctatactttaatllaagaacaa	1941
Db	4379	CTTTTGTTGTTATTTCTGTGTTTAATTTTTTTTGTATCAATCTTAAATTTAGTTTAA	4438
QY	1942	attcttgaaaaatgltataatcttcaacaatlataaatalatgylgcctaaatglat	2001
Db	4439	ATTTTATATATATTTCTAATTTTAAATTTAATTTAATTTGTTGAATTTAAATTTATTAAT	4498
QY	2002	tctctatgctcttaaaaatatlltlltctatlltagtctaaatcatatgaaccaata	2061
Db	4499	TTCTAATTTTTTTTGAACAACACCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAA	4558
QY	2062	atagctggtgaatcaaatatcccatctaatatcttcttgaaaatcacaatatatcat	2121
Db	4559	TTTGTTTAAATTTGTTTACCATTTTTTTTTTTAACTCATTTTACCCTATTTAAATTTATTTA	4618
QY	2122	ttagcaatatacaatgcatagaaagttcccaaaaaaatcttglttaacagaactctcaaa	2181
Db	4619	TTATTTATCGAAAAAAACCTTAAAAAAGCACTTGTGTAATTTTATTTGTATTAATAACGAAAA	4678
QY	2182	ttcttcttcttcttaatggaacagaataacagatagaaaaactatctgtgtggaatgaa	2241
Db	4679	TGACATATATATAAAGAAAAAAATTAATAAGAAATTAATAAAAAACAAGGGTTGATTACTTT	4738
QY	2242	gtagtaatatagcttaagcaaatcttaaaaaattataag 2282	
Db	4739	TAACTAAGTCAAGTCTAATAATTAATTTCTATTTTTTTTTTAAAG 4779	

Query Match	Best Local Similarity	0.9%;	Score 62.6;	DB 1.	Length 7218;
Matches 26;	Conservative 151;	Mismatches 90;	Indels 0;	Gaps 0;	
QY	4100	actcattataggttttaattatttatttattg	ttgttaataa	atc	aatg
Db	1486	actgtaattactactctatctgcagtaggtt	aaagagata	gaa	aaatttggtac
QY	4160	tgtatcagcagcaacgatatgtaagttcat	tgtagag	gaaatg	gagagaaatgg
Db	1426	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr
QY	4220	atatacgaagaagaatcaagaagaagaag	tgaaaac	agagat	gaagcagagatga
Db	1366	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr
QY	4280	gtaagagatcatatgtagaacaacaagaac	agatgatc	ctt	ggatgg
Db	1306	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr
QY	4340	aaacatggaattatcagcagagcaaa	4366		
Db	1246	rrrrrrrrrrrrrrrrrrrrrrrrrrrrrr	1220		

[illegible]

Db 5607 AAAAAAAAAAATGTAATTTGGAATTAATAAATGTAAGGCTTTTTTTT 5548
QY 2007 atgtcttaaaatatttttttata-----ttaagtaaaatcatatgaacca 2059
Db 5547 AAAATTAATGATGATTTTATTTTAAATGATGACAGATTAATAATCTTAAACAA 5488
QY 2060 taatagtcgtggaattcaaatcatctcatatatttttgaatctcaaatattaat 2119
Db 5487 TAAACATATTTGATTTTTTTTTTTTTTTTTTTTTTCAATTAATTAATA 5428
QY 2120 attagtaacaacatcatagaaagttccaaaaaatttgtaacagaaactcca 2179
Db 5427 TAAATATTAATATCTTGATGAACTTCAATTTTATTAATTAATTTTAA 5368
QY 2180 aattttttttt 2192
Db 5367 AATATGATGATCT 5355

RESULT 13
US-08-817-926-27/c
; Sequence 27, Application US/08817926
; Patent No. 6001590
; GENERAL INFORMATION:
; APPLICANT: komeda, Toshihiro
; APPLICANT: Suda, Hisako
; APPLICANT: Tamai, Yukio
; APPLICANT: Iwamatsu, Akihiro
; APPLICANT: Kato, No. 6001590uo
; APPLICANT: Sakai, Yasuyoshi
; TITLE OF INVENTION: PROMOTER/TERMINATOR FOR CANDIDA BOIDINI
; TITLE OF INVENTION: FORMATE DEHYDROGENASE GENE
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,926
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/02597
; FILING DATE: 12-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 234133/1995
; FILING DATE: 12-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 42536/1996
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 29,768
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4818 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Candida boidini
US-08-817-926-27

Query Match 0.98; Score 60.8; DB 3; Length 4818;
Best Local Similarity 46.3%; Pred. No. 0.0033;
Matches 275; Conservative 0; Mismatches 312; Indels 7; Gaps 2;

QY 54 ttcggttatctcgttcacatgattgagtttggttcacatcagattccagaataat 113
Db 4718 TTTTATTTATTTCTTTTATTTTATTTTCTTATATATGATTTTGGCTTATTAACAATA 4659
QY 114 aaatcaaaatcattcaaaatattcaagaatcaatcaatcattcattcattgatt 173
Db 4658 AAATTTAGAGTCCCTTTTAAAGTTTTCGATTAATTAATTAATTAATTAATAGGT 4599
QY 174 atacaacaatctcatcttggtatagagaataataggagttcggaattataata 233
Db 4598 AAATAGATTAATAAATAAATGATACCAATTTAAACAAATTTAATTAATTAATAATA 4539
QY 234 a-aggaataatcgattccattggttggtatcacagttgaatttctggtttctt 292
Db 4538 TTCTAAATTTTAAACGTTTGTTCAAAAAATTAGAAATTAATTAATTAATTAATTA 4479
QY 293 ttatagtataatgaatcaaaagagatgattgagagtgtaaacataattcgatt 352
Db 4478 CAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAGAT 4419
QY 353 gaccccaaaaaaacaacaacaacaacaccccccccgatataagtttgggt 412
Db 4418 TGATACAAAAAATAAATAAACAAGAAATTAACAACAAACAAACAAATTTTCTTTT 4359
QY 413 tctgattggttatttgatca-----taattacgacatcttcttgattactat 466
Db 4358 TTTTATTTACCAATTAATTAACCAATTAACCAATTAATTAATTAATTAATTAATTA 4239
QY 467 gaagatttcttacaatcaaatcaaatcgaattcatctcttgattataatcaat 526
Db 4298 AAAGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATAC 4239
QY 527 gagtgtaataatccgttatccgattcgaatccatcatgattatgatttcttgccaatca 586
Db 4238 TACTCTCTTTTAAACGTTTGACGTAATTAATTAATTAATTAATTAATTAATTAATCT 4179
QY 587 gcaattattacaagagattgagaacaaacgaaataagaaaggaaga 640
Db 4178 TTCAATTTCAATTTTAAATTTGAATAAACAAGGAAATTAATAAAGAGGAAA 4125

RESULT 14
US-08-998-416-288/c
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPYII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NO. 6239264ch Carolina
; COUNTRY: USA
; ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 288:
SEQUENCE CHARACTERISTICS:
LENGTH: 837 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1241RP
US-08-998-416-288

Query Match 0.9%; Score 60.6; DB 4; Length 837;
Best Local Similarity 49.1%; Pred. No. 0.0023;
Matches 191; Conservative 0; Mismatches 194; Indels 4; Gaps 1;

QY 1939 aaaaatttggaaaatgataattcttaacaatatatttaaatatgagtcctataatg 1998
DB 608 AATAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 549
QY 1999 tatttcctagtccttaaaatattttttattattagtttaataacatcatgaca 2058
DB 548 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 489
QY 2059 ataatagttgatgaatcaaatatccatcaatatttttgaatctcaaatattaa 2118
DB 488 GAATTTAAAGTTAAATTAATTTTAAATTAATTTCTTATAAAGCTTAAATAATTA 429
QY 2119 tatttagtaataacaatgcatagaagttccaaaaaaatttgttaacagaacttcc 2178
DB 428 TCAACCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 369
QY 2179 aaaaatttcttttaagacaagaataacagatagaactatttgggtggaatg 2238
DB 368 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 313
QY 2239 gaagtagtaataacaatgaagaatttaaaaaattatataagcctacagcgctcaa 2298
DB 312 AATATATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 253
QY 2299 gtagtctacgtgagtgtaattataa 2327
DB 252 AATATTTTAACTAATTAATTAATTAATTTGAA 224

RESULT 15
US-08-446-855A-1/C
Sequence 1, Application US/08446855A
Patent No. 5849573
GENERAL INFORMATION:
APPLICANT: Stewart, Thomas S
APPLICANT: Flores, Maria V
APPLICANT: O'Sullivan, William J
TITLE OF INVENTION: Nucleotide sequence encoding cardamyl

TITLE OF INVENTION: phosphate synthetase II
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye PC
STREET: 1100 NO. 5849573th Glebe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 47-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic
US-08-446-855A-1

Query Match 0.9%; Score 60.4; DB 2; Length 8920;
Best Local Similarity 48.8%; Pred. No. 0.0047;
Matches 163; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 1919 ttacatttaatttaaaagaacaattttgaaaatgtataattcttaacaattatt 1978
DB 8751 TTTTATTTTATTAATTAATTAATTAATTTTAAATTAATTAATTAATTTGTAATTTGAA 8692
QY 1979 aaaaatgagtcctaaatgataatttccatagttcttaaaatatttttttattattg 2038
DB 8691 AAACCATTTTGGTTATACATATGAGTGAATTAATTAATTTGTAATTAATTAATTAAT 8632
QY 2039 tataatacatatgaaccaataatgattggtgaattcaaatatattccattatattt 2098
DB 8631 TATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8572
QY 2099 tgaattctacaattatataatttagtcaatcaacaatgcatagaagttccaaaaaa 2158
DB 8571 TTAACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8512
QY 2159 ttctgttaacagaacttcaaatatttttttttaaggaacaagaataacagataga 2218
DB 8511 TATTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8452
QY 2219 aactatttgcgtggaatggaatgtaataa 2252
DB 8451 AATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8418

Search completed: September 29, 2001, 11:27:42
Job time: 16641 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2001, 07:23:36 ; Search time 378.88 seconds
(without alignments)
11415.168 Million cell updates/sec

Title: US-09-502-426-1

Perfect score: 6888
Sequence: 1 atcgtggtattatattgttgccttagtcagccaagcctt 6888

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: N_Geneseq_0601:*
2: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:*
23: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6888	100.0	6888	21	AAA59599
2	251	3.6	936	22	AAF58252
3	251	3.6	936	22	AAF58254
4	251	3.6	936	22	AAF58257
5	251	3.6	936	22	AAF58259
6	251	3.6	936	22	AAF58262
7	251	3.6	936	22	AAF58255
8	250.6	3.6	936	22	AAF58252
9	250.6	3.6	936	22	AAF58254
10	250.6	3.6	936	22	AAF58257
11	250.6	3.6	936	22	AAF58259

12	250.6	3.6	936	22	AAF58262	Oligonucleotide D2
13	250.6	3.6	936	22	AAF58255	Oligonucleotide D1
14	92.8	1.3	4937	18	AAAT85307	Arabidopsis thalia
15	88.2	1.3	244	22	AAF58238	Oligonucleotide D1
16	87.6	1.3	1608	18	AAAT85306	Arabidopsis thalia
17	87.6	1.3	1646	21	AAAT85307	Arabidopsis thalia
18	87.6	1.3	1649	21	AAAT85306	Arabidopsis thalia
19	85.4	1.2	244	22	AAF58238	Oligonucleotide D1
20	79.4	1.2	8310	20	AAAT29911	CDNA encoding a SC
21	77.4	1.1	19124	18	AAAT2882	Plasmodium var-7 g
22	77.4	1.1	19124	21	AAAT2882	Plasmodium var-7 p
23	74	1.1	2104	13	AAAT25273	Sequence encoding
24	73	1.1	1864	8	AAAT71405	Sequence of ANS-1
25	70.8	1.0	8310	20	AAAT29911	CDNA encoding a SC
26	69.6	1.0	4590	7	AAAT6472	Sequence encoding
27	68.8	1.0	875	21	AAAT1920	Human colon cancer
28	68.6	1.0	1864	8	AAAT71405	Sequence of ANS-1
29	68.2	1.0	2104	13	AAAT25273	Sequence encoding
30	67.4	1.0	5760	6	AAAT50530	Sequence encoding
31	67.4	1.0	19124	18	AAAT2882	Plasmodium var-7 g
32	67.4	1.0	19124	21	AAAT2882	Plasmodium var-7 p
33	67.2	1.0	875	21	AAAT1920	Human colon cancer
34	66.6	1.0	1398	21	AAAT8714	Arabidopsis thalia
35	66	1.0	2486	21	AAAT7037	Nucleotide sequenc
36	65.8	1.0	2486	21	AAAT7037	Nucleotide sequenc
37	65.4	0.9	2503	15	AAAT53480	PNPX30 xylanase cd
38	65.4	0.9	20674	21	AAAT58017	Arachidonic acid m
39	65.2	0.9	6033	21	AAAT0152	Plasmodium falcipa
40	65	0.9	2418	13	AAAT27886	P.falciparum GBP3
41	64.4	0.9	5409	21	AAAT0151	Plasmodium falcipa
42	64.2	0.9	1431	21	AAAT37082	DNA sequence encod
43	64.2	0.9	1671	13	AAAT24134	50 kD subunit of S
44	63.6	0.9	8920	15	AAAT62924	Carbamoyl-phosphat
45	63	0.9	20674	21	AAAT58017	Arachidonic acid m

ALIGNMENTS

RESULT 1	
ID	AAA59599 standard; DNA; 6888 BP.
AA59599	
AC	AAA59599;
XX	
DT	14-NOV-2000 (first entry)
XX	
DE	DNA encoding a cytochrome P450 enzyme designated DMF4.
XX	
KW	DMF4; cytochrome P450 enzyme; brassinosteroid; 22alpha-hydroxylation;
KW	plant phenotype; cell elongation; ss.
XX	
OS	Arabidopsis sp.
XX	
XX	
FT	Key
FT	Location/Qualifiers
FT	1..3203
FT	promoter
FT	/*tag= a
FT	3060..3125
FT	TATA_signal
FT	/*tag= b
FT	3203..6110
FT	CDS
FT	/*tag= c
FT	/*product= "DMF4"
FT	/*note= "contains introns"
FT	3203..3423
FT	exon
FT	/*tag= d
FT	3424..3503
FT	intron
FT	/*tag= e
FT	3504..3828
FT	exon
FT	/*tag= f
FT	3829..3913
FT	intron
FT	/*tag= g
FT	3914..4066
FT	exon
FT	/*tag= h

QY	1261	gggatgctatacccttcacgctatgctactagagactaaagaatagtatgctgctc	1320
Db	1261	gggatgctatacccttcacgctatgctactagagactaaagaatagtatgctgctc	1320
QY	1321	gataaatgnaatlcacacgctggtglaataatlatggaacgctatgctcgcctcgaa	1380
Db	1321	gataaatgnaatlcacacgctggtglaataatlatggaacgctatgctcgcctcgaa	1380
QY	1381	atatactctctggttgcacacaataaaaaaacaagaagaaaaacgaattctt	1440
Db	1381	atatactctctggttgcacacaataaaaaaacaagaagaaaaacgaattctt	1440
QY	1441	cttgatcccttcacatgatacctaataatgcatagatccttgggttacaggttcgaagtc	1500
Db	1441	cttgatcccttcacatgatacctaataatgcatagatccttgggttacaggttcgaagtc	1500
QY	1501	tcataacgctgttaacacacctgcgcaacttaaatgtcttctttaaigtacatcttaaat	1560
Db	1501	tcataacgctgttaacacacctgcgcaacttaaatgtcttctttaaigtacatcttaaat	1560
QY	1561	attatgtgtgttggaatttaataagagcgaactgttaacattacaatatattatag	1620
Db	1561	attatgtgtgttggaatttaataagagcgaactgttaacattacaatatattatag	1620
QY	1621	atactagtatgtatattccaatacacttcttgatggtttaaacttaactgttct	1680
Db	1621	atactagtatgtatattccaatacacttcttgatggtttaaacttaactgttct	1680
QY	1681	tcctacggtataaatatataatcatcgcgaagttaaaaaagtttgccttaatttcgcgcatgc	1740
Db	1681	tcctacggtataaatatataatcatcgcgaagttaaaaaagtttgccttaatttcgcgcatgc	1740
QY	1741	atgagagtaaaccttaagacttaatttttggaaaagttaacccctttactacatagatt	1800
Db	1741	atgagagtaaaccttaagacttaatttttggaaaagttaacccctttactacatagatt	1800
QY	1801	aatacgcgtatgtttgtctgcacataagacagccttcacacgttgcatagtcaattct	1860
Db	1801	aatacgcgtatgtttgtctgcacataagacagccttcacacgttgcatagtcaattct	1860
QY	1861	tcctgcaaatatnaaatltagaattcaatgcactacataagaagaacagctgctgatt	1920
Db	1861	tcctgcaaatatnaaatltagaattcaatgcactacataagaagaacagctgctgatt	1920
QY	1921	acattttaattaaagacaaaatttttggaaaagttaataatttccaataattattaa	1980
Db	1921	acattttaattaaagacaaaatttttggaaaagttaataatttccaataattattaa	1980
QY	1981	aataatgctctataatglatctcctatgttcttaaaaaatattttttatatactagta	2040
Db	1981	aataatgctctataatglatctcctatgttcttaaaaaatattttttatatactagta	2040
QY	2041	taataacattatgaacaaataatagtctggtgaattccaatatctccattataattttg	2100
Db	2041	taataacattatgaacaaataatagtctggtgaattccaatatctccattataattttg	2100
QY	2101	aaatctaaatattatattattagtcataaacaatgcataagaaagtcccaaaaaaat	2160
Db	2101	aaatctaaatattatattattagtcataaacaatgcataagaaagtcccaaaaaaat	2160
QY	2161	ttgttaacagaactccaatattttttttaaagacaagaataaacaagatagaataa	2220
Db	2161	ttgttaacagaactccaatattttttttaaagacaagaataaacaagatagaataa	2220
QY	2221	ctattttgtgtggaatggaatgataataatcaatgaacaaattttaaaaaattatata	2280
Db	2221	ctattttgtgtggaatggaatgataataatcaatgaacaaattttaaaaaattatata	2280
QY	2281	agccatacgcgctcaaaagtatgtatctagtaggtgaatnaaatgaatgctgcat	2340
Db	2281	agccatacgcgctcaaaagtatgtatctagtaggtgaatnaaatgaatgctgcat	2340
QY	2341	tcagaattgggacaacaatgaaaaagaaatlaaaatatttaactttaaaataaataa	2400
Db	2341	tcagaattgggacaacaatgaaaaagaaatlaaaatatttaactttaaaataaataa	2400
QY	2401	ttgagtaaatgttcttctgactatlgaggcgcaaaaaaagacaaatgccaagaatctac	2460
Db	2401	ttgagtaaatgttcttctgactatlgaggcgcaaaaaaagacaaatgccaagaatctac	2460
QY	2461	gggttgactctgcagttccgttaataacttaataacttgccttgaccgcgaacgtctg	2520
Db	2461	gggttgactctgcagttccgttaataacttaataacttgccttgaccgcgaacgtctg	2520
QY	2521	taagggtccttcgacatttcaactgttccaccccttaactggtgagccacccctttcca	2580
Db	2521	taagggtccttcgacatttcaactgttccaccccttaactggtgagccacccctttcca	2580
QY	2581	latcctaaggatlaatttggaatcccaatttaaacgcatlgagacgctacccgacttcc	2640
Db	2581	latcctaaggatlaatttggaatcccaatttaaacgcatlgagacgctacccgacttcc	2640
QY	2641	tggaattctgtcggagcatltaaaaaatratagcagaaatgggttatattatataa	2700
Db	2641	tggaattctgtcggagcatltaaaaaatratagcagaaatgggttatattatataa	2700
QY	2701	aactcaacacttgatcagataaaaatttcaataacacctttacgcatgcttcgacgatct	2760
Db	2701	aactcaacacttgatcagataaaaatttcaataacacctttacgcatgcttcgacgatct	2760
QY	2761	atctaatgacttttttttcttcaacacggtggaatgaatagatctatagccagag	2820
Db	2761	atctaatgacttttttttcttcaacacggtggaatgaatagatctatagccagag	2820
QY	2821	acaattgttatagatataatccatgaattatgaattatgaattatgaattatgaattat	2880
Db	2821	acaattgttatagatataatccatgaattatgaattatgaattatgaattatgaattat	2880
QY	2881	ctatttcagcatcgcagatttctgcacatttgcgaatttgcgaatttgcgaatttgcga	2940
Db	2881	ctatttcagcatcgcagatttctgcacatttgcgaatttgcgaatttgcgaatttgcga	2940
QY	2941	aagtaataaaggagacataacgaggaacaaagaatgaacacgagagaacaaagcca	3000
Db	2941	aagtaataaaggagacataacgaggaacaaagaatgaacacgagagaacaaagcca	3000
QY	3001	tgaagctcatgctggttagttgaagcttaagaagaatttataaattttaatgacgtg	3060
Db	3001	tgaagctcatgctggttagttgaagcttaagaagaatttataaattttaatgacgtg	3060
QY	3061	ataacaattataatttctgcgacttcttlaaaacccctcttaacaacagagctcccttt	3120
Db	3061	ataacaattataatttctgcgacttcttlaaaacccctcttaacaacagagctcccttt	3120
QY	3121	tcagtaagaagctcgcattcccaatcttaaaagacaaagccattagaagaagatgagta	3180
Db	3121	tcagtaagaagctcgcattcccaatcttaaaagacaaagccattagaagaagatgagta	3180
QY	3181	gagagagagagaactagctccatgttgcgaacagagacataactcttcccttcc	3240
Db	3181	gagagagagagaactagctccatgttgcgaacagagacataactcttcccttcc	3240
QY	3241	tccttcccatcgccttgcctctctctctctctcttcttgaattctctgaaagagaataag	3300
Db	3241	tccttcccatcgccttgcctctctctctctctcttcttgaattctctgaaagagaataag	3300
QY	3301	aaaaacagatccaatccaactcccggtgtaaaatccggttgcgaattcttggtaaacat	3360
Db	3301	aaaaacagatccaatccaactcccggtgtaaaatccggttgcgaattcttggtaaacat	3360
QY	3361	cggtatctttaaaccgttaacacgcgaacaacactggttactcttgaacaaacatgcttc	3420
Db	3361	cggtatctttaaaccgttaacacgcgaacaacactggttactcttgaacaaacatgcttc	3420
QY	3421	caagtaaacacaacatcttccaaaactcaaaaaataaatcctctgtttttgaaatt	3480
Db	3421	caagtaaacacaacatcttccaaaactcaaaaaataaatcctctgtttttgaaatt	3480

Db 3421 caagtaacaacaacatctccaaaaaactcaaaaaataaactcctgcttttggaaattc 3480
QY 3481 gactaaatgtgttaattttaaaggtaatgtaagataatagatcgaactgtttggaa 3540
Db 3481 gactaaatgtgttaattttaaaggtaatgtaagataatagatcgaactgtttggaa 3540
QY 3541 ccaacgaatcgtatcgaactgtgtgacttaataatagatcattacaacaaaggaagg 3600
Db 3541 ccaacgaatcgtatcgaactgtgtgacttaataatagatcattacaacaaaggaagg 3600
QY 3601 cctcttaaatgtatgttaccatagaagataggtgggaattcttggaatagtcgaatc 3660
Db 3601 cctcttaaatgtatgttaccatagaagataggtgggaattcttggaatagtcgaatc 3660
QY 3661 gtctctgttggtgacatgcaatgcaatagatagatgcaatgcaatgcaatgcaatgca 3720
Db 3661 gtctctgttggtgacatgcaatgcaatagatagatgcaatgcaatgcaatgcaatgca 3720
QY 3721 gcaacgtcttgaacatcttcaacttaaaagatgtgagagacatctgttctgac 3780
Db 3721 gcaacgtcttgaacatcttcaacttaaaagatgtgagagacatctgttctgac 3780
QY 3781 tcttggaacaaacatctatcttctgctcaagacgaagccaaaggttttaattt 3840
Db 3781 tcttggaacaaacatctatcttctgctcaagacgaagccaaaggttttaattt 3840
QY 3841 atcttatttctgctaatttcttctgttataatccttgaagtttcaacttttct 3900
Db 3841 atcttatttctgctaatttcttctgttataatccttgaagtttcaacttttct 3900
QY 3901 ttttaattgaaacatttgaactttaaactaaatgagaaagataatgagttgac 3960
Db 3901 ttttaattgaaacatttgaactttaaactaaatgagaaagataatgagttgac 3960
QY 3961 agagaagaagaacgaacatcttaagaagaagatgtaacttcaatgaaagattgctc 4020
Db 3961 agagaagaagaacgaacatcttaagaagaagatgtaacttcaatgaaagattgctc 4020
QY 4021 tgcctccttaacttaccagaactgttataccttaagccttcaagttacattttt 4080
Db 4021 tgcctccttaacttaccagaactgttataccttaagccttcaagttacattttt 4080
QY 4081 ttttgctgtaagatcgaactctcatatagtttttaattttatttttttggttaact 4140
Db 4081 ttttgctgtaagatcgaactctcatatagtttttaattttatttttttggttaact 4140
QY 4141 aaaaatcctaaatgtgtgtgtgtagtcaagacgaagataatgaaatgaggaagaa 4200
Db 4141 aaaaatcctaaatgtgtgtgtgtagtcaagacgaagataatgaaatgaggaagaa 4200
QY 4201 atggaagaagaagaatctgatacacaagaagaatcacaagaagaagaatgaaaca 4260
Db 4201 atggaagaagaagaatctgatacacaagaagaatcacaagaagaagaatgaaaca 4260
QY 4261 gaagataaagacagagatgataagagatgataagatgataagacaaagacgaatgac 4320
Db 4261 gaagataaagacagagatgataagagatgataagatgataagacaaagacgaatgac 4320
QY 4321 cttttggatgggttttgaacacatcgaatttcaacgaagacaaatctcgaatc 4380
Db 4321 cttttggatgggttttgaacacatcgaatttcaacgaagacaaatctcgaatc 4380
QY 4381 ctttaagtttgaatttgcgcgaacatgagacttctctgtagacatctgctcgaatc 4440
Db 4381 ctttaagtttgaatttgcgcgaacatgagacttctctgtagacatctgctcgaatc 4440
QY 4441 tttctggaacgttgcgcgaacatgagacttctctgtagacatctgctcgaatc 4500
Db 4441 tttctggaacgttgcgcgaacatgagacttctctgtagacatctgctcgaatc 4500
QY 4501 aagtttaattcacaacatgtaagatataaagtttataagattatctatag 4560
Db 4501 aagtttaattcacaacatgtaagatataaagtttataagattatctatag 4560

QY 4561 aatatacgaatgaaaaaagatgataatgataatgataatgataatgataatgataatgata 4620
Db 4561 aatatacgaatgaaaaaagatgataatgataatgataatgataatgataatgataatgata 4620
QY 4621 tactttggaaggaagacatcttgagatcgcgagggcgcaagaaggaactgagagtcag 4680
Db 4621 tactttggaaggaagacatcttgagatcgcgagggcgcaagaaggaactgagagtcag 4680
QY 4681 aatataatggaatgataaagaagaatggaacttacttacttacttacttacttacttact 4740
Db 4681 aatataatggaatgataaagaagaatggaacttacttacttacttacttacttacttact 4740
QY 4741 ctcattatattctatgltcatalatgattatgataagaaacaaatbatgtattttt 4800
Db 4741 ctcattatattctatgltcatalatgattatgataagaaacaaatbatgtattttt 4800
QY 4801 ttggtgtgtgtgaaggtataaataatgaaactctcgatgtggaaatgagttagttttg 4860
Db 4801 ttggtgtgtgtgaaggtataaataatgaaactctcgatgtggaaatgagttagttttg 4860
QY 4861 catcgcaaaacactcacaagaatgttccgtacaaagtaaaacttactacaaatttta 4920
Db 4861 catcgcaaaacactcacaagaatgttccgtacaaagtaaaacttactacaaatttta 4920
QY 4921 aataatgaaatccggaatattgaaatccttattgataagaaatataaataatattacat 4980
Db 4921 aataatgaaatccggaatattgaaatccttattgataagaaatataaataatattacat 4980
QY 4981 tttcttaattgttggaaaaaaagatacgaatcctctagttgggttgaaggtgttaccggtgac 5040
Db 4981 tttcttaattgttggaaaaaaagatacgaatcctctagttgggttgaaggtgttaccggtgac 5040
QY 5041 ctacgcgtatcatcttgatataatctcgtatacgaacaaactccttcaatccttggag 5100
Db 5041 ctacgcgtatcatcttgatataatctcgtatacgaacaaactccttcaatccttggag 5100
QY 5101 atggcaacgaatgataaagaatctcctcgttcaactatcgaaatagtgatagttt 5160
Db 5101 atggcaacgaatgataaagaatctcctcgttcaactatcgaaatagtgatagttt 5160
QY 5161 ttatcattatgcatgaaatgataatgataatgataatgataatgataatgataatgata 5220
Db 5161 ttatcattatgcatgaaatgataatgataatgataatgataatgataatgataatgata 5220
QY 5221 agaaacttaacttgcgcgaacaaacttcttataatcttataatcttcaagttagatcttaaca 5280
Db 5221 agaaacttaacttgcgcgaacaaacttcttataatcttataatcttcaagttagatcttaaca 5280
QY 5281 cgtcattgataatctgatacacaatgaaagacataaataaataaagaagaagaagac 5340
Db 5281 cgtcattgataatctgatacacaatgaaagacataaataaataaagaagaagaagac 5340
QY 5341 gtgtgtgtgaaatgtgacgtggaagacaaatattagtagtagtagtagtagtagtagtag 5400
Db 5341 gtgtgtgtgaaatgtgacgtggaagacaaatattagtagtagtagtagtagtagtagtag 5400
QY 5401 taggagctgcctaaatattcttataatcgtgacgttgaatttataatgtagtagtag 5460
Db 5401 taggagctgcctaaatattcttataatcgtgacgttgaatttataatgtagtagtag 5460
QY 5461 gatacaatttaagaatctgaaaggaagatgatacgaagcttgcgaatctcacaacgccc 5520
Db 5461 gatacaatttaagaatctgaaaggaagatgatacgaagcttgcgaatctcacaacgccc 5520
QY 5521 atgactgtgacgttgcgttagatctgcgaagcttgcgaagcttgcgaagcttgcgaagct 5580
Db 5521 atgactgtgacgttgcgttagatctgcgaagcttgcgaagcttgcgaagcttgcgaagct 5580
QY 5581 cttatattatataaacaacatacaaatcggcggaattagagaataataataataatagata 5640
Db 5581 cttatattatataaacaacatacaaatcggcggaattagagaataataataataatagata 5640

QY	5641	tttagattattatagatatataactcttattgaaataagaataaataagtgatcatgatgttta	5700
Db	5641	ttatgatattactatgagatatataactcttaagaaataagaataaataagtgatcatgatgttta	5700
QY	5701	taqtgagttggggaaggttaagaagtggtgtgagaaatataatgaacccacggtgtgtgtg	5760
Db	5701	taqtgagttggggaaggttaagaaggtgtgtgagaaatataatgaacccacggtgtgtgtg	5760
QY	5761	ccaacaagacggtgtctctctctctctcttcttctcccaactctcttctttgtggggtttatc	5820
Db	5761	ccaacaagacggtgtctctctctctctcttcttctcccaactctcttctttgtggggtttatc	5820
QY	5821	gtgattatataaactcgtgtttgcgtttttttttttgtgcagagcagaacaacgagacgt	5880
Db	5821	gtgattatataaactcgtgtttgttcgtttttttttgtgcagagcagaacaacgagacgt	5880
QY	5881	catcgtcaagaagtgtagtttttcgcacgttgagggaacaactacaactcgcgttttgagag	5940
Db	5881	catcgtcaagaagtgtagtttttcgcacgttgagggaacaacaactacacgtcgttttgagag	5940
QY	5941	ggccaagagcctatgtgtgtgttcgaagcttagccaagttaagaattgagcagtggtttatc	6000
Db	5941	ggccaagagcctatgtgtgtgttcgaagcttagccaagtttagaattgagcagtggtttatc	6000
QY	6001	atctagtctcttaaatccaattggaattggaattgagcaagaatgataccaacttgctttccct	6060
Db	6001	atctagtctcttaaatccaattggaattgagcaagaatgataccaactatgctttccct	6060
QY	6061	ttgtgtgtttctcctaacggtttgcacctatagaggtttccgcatacttcgtataaataaaaaa	6120
Db	6061	ttgtgtgtttctcctaacggtttgcacctatagaggtttccgcatacttcgtataaataaaaaa	6120
QY	6121	agatgaaagatattttatctctctctcttttttttttgataaatttaaatcatcttttttgc	6180
Db	6121	agatgaaagatattttatctctctctcttttttttttgataaatttaaatcatcttttttgc	6180
QY	6181	ccaatgatatataaanaatttgatataataatattattgtgatatcggtttttttagttgcgg	6240
Db	6181	ccaatgatatataaanaatttgatataataatattattgtgatatcggtttttttagttgcgg	6240
QY	6241	tttagaanaaggggttcgcacttttcgaaatggaacatgatatataatgtggagagcagaagt	6300
Db	6241	tttagaanaaggggttcgcacttttcgaaatggaacatgatatataatgtggagagcagaagt	6300
QY	6301	gagcgttttgacactttgtatgtgattgttgbattatagttgtgcacactttaacact	6360
Db	6301	gagcgttttgacactttgtatgtgattgttgbattatagttgtgcacactttaacact	6360
QY	6361	aaatgagccttctataagcccaattatatttcgattataaacaagtgacaactttact	6420
Db	6361	aaatgagccttctataagcccaattatatttcgattataaacaagtgacaactttact	6420
QY	6421	tcggtttttgttcggaagcaataaactgttcaaatcaccaacagaagaattatgttaac	6480
Db	6421	tcggtttttgttcggaagcaataaactgttcaaatcaccaacagaagaattatgttaac	6480
QY	6481	actcgtgtgtcctagtgagggaataactttggccttgagagctgacatacgaacaagaagaa	6540
Db	6481	actcgtgtgtcctagtgagggaataactttggccttgagagctgacatacgaacaagaagaa	6540
QY	6541	ccgtgtcaatatgagatacacctctctatgacggtttcccaacttlatcttgattctatcag	6600
Db	6541	ccgtgtcaatatgagatacacctctctatgacggtttcccaacttlatcttgattctatcag	6600
QY	6601	tgtagaactgacacaagaagtttagtggttcaaaaagactaaatagaataacaatagctctca	6660
Db	6601	tgtagaactgacacaagaagtttagtggttcaaaaagactaaatagaataacaatagctctca	6660
QY	6661	ggataaagaagtttcaataaagttgttcttattttgagaagaagaagagagactttta	6720
Db	6661	ggataaagaagtttcaataaagttgttcttattttgagaagaagaagagagagactttta	6720
QY	6721	ctgttctctgtgtctatctatctctttaaagagaggttgcgttttacttctctatctca	6780

[illegible]

[illegible]

RESULT	4
AAF58257/c	
ID	AAF58257 standard; DNA; 936 BP.

AC AAF58257;

DT 24-APR-2001 (first entry)

DE₁ Oligonucleotide D1954.

KW Electron-transfer group; **ETM**; mismatch; genotyping; zone averaging; **ss**

Crustaceans

XX
XX
MO300107666-83

XX
DD 01-FEB-2001XX
XX
26-THR-2000

XX	36-TWT-1000.	00HC-014EC0E
DB		

PR 17-MAR-2000; 2000US-0190259
XX

PA (CLIN-) CLINICAL MICRO SENSORS

PI	Umek RM;
yy	

DR WPI; 2001-159728/16
xx

PT Nucleic acids conta

PT a single surface -

PS Example 6; Page 127; 159pp; English.

CC The present invention relates to a composition comprising two nucleotide
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.

SQ Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other.

Query Match	3.6%	Score 251	DB 22	Length 936
Best Local Similarity	1.0%	Pred. NO. 1.2e-30		
Matches	8	Conservative	184	Indels 0
		Mismatches	184	Gaps 0

QY 1508 ggcgtataacatctgcgaacattaaatgtcttctttaatgacatctttaacattattt 1567
 Db 793 GGGGGCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 734
 QY 1568 gttagtggaaatttaagaagcgaaactgttaacatlaaatattatagatactag 1627
 Db 733 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 674
 QY 1628 taatgattatccaaataacataacttggattgaactaaactgtttcttccatag 1687
 Db 673 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 614
 QY 1688 gtataaatataatcatcgcagggttaaaaaagtttcttattttcgcgagtcagaaag 1747
 Db 613 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 554

QY	1748	ataaacctaagcttaattttttgaaatgtaaacctttaactatcatagatattaac	1807
Db	553	wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww	494
QY	1808	gtatgttttctgcataatgtaagcctctacaactgtaagtcattttcttgca	1867
Db	493	wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww	434
QY	1868	atattaattagaattcaatgctactatcatagagaagaacagctgattacattt	1927
Db	433	wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww	374
QY	1928	aatttaagaacaatttttgaaaaatgatatattttccaacatatcttaaatatga	1987
Db	373	wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww	314
QY	1988	tgccataatgctattctcctatgctcttaaatatttttttattatgttataatc	2047
Db	313	wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww	254
QY	2048	attatgaaaccaataatgttggtaattcaataatcctcatatatttttgaaatcta	2107
Db	253	wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww	194
QY	2108	caaatatataatttttgtaacatacatgcatagaaagtccaacaaaatttgytaa	2167
Db	193	wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww	134
QY	2168	cagaacctccaattttttttttagaacaagaataacagatagaacattttt	2227
Db	133	wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww	74
QY	2228	gttgygaatggaagtagtaataacataaagcaaaattttaaataatataaagctat	2287
Db	73	wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww	14
QY	2288	a 2288	
Db	13	w 13	

```

RESULT 5
AAF58259/C
ID AAF58259 standard; DNA. 936 BP.
XX
AC AAF58259;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D2004.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
   gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
XX
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI
XX
PI Umek RM;
XX
DR
XX
WI: 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses or

```


XX	AAF58252;
AC	
XX	24-APR-2001 (first entry)
DT	
XX	Oligonucleotide D1835.
DE	
XX	Electron-transfer group; ETW; mismatch; genotyping;
KW	gene expression; ss.
XX	
XX	Synthetic.
OS	
PN	WO200107665-A2.
XX	
PD	01-FEB-2001.
XX	
PF	26-JUL-2000; 2000MO-US20476.
XX	
PR	26-JUL-1999; 99US-0145695.
PR	17-MAR-2000; 2000US-0190259.
XX	
PA	(CLIN-) CLINICAL MICRO SENSORS INC.
XX	
P1	Imek RM;
XX	
DR	WPI; 2001-159728/16.
XX	
XX	
PT	Nucleic acids containing electron-transfer group, useful as labels in
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT	a single surface -
XX	
PS	Example 6; Page 127; 159pp; English.
XX	
CC	The present invention relates to a composition comprising two nucleic
CC	acids each containing an electron-transfer group (ETW) having
CC	different redox potentials. The invention is used for electronic
CC	detection of nucleic acids, especially of substitutions (mismatches)
CC	and single-nucleotide polymorphisms, e.g. for genotyping,
CC	monitoring gene expression.
XX	
50	Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

```

Query Match          3.6%; Score 250.6; DB 22; Length 936;
Best Local Similarity 0.58; Pred. No. 1.4e-30;
Matches      4; Conservative 596; Mismatches 185; Indels       0; Gaps
OY 1544 taatgcaccttaacatatattctgttgtaacttaataagaagcaactgtaaacat 16030
Db 1 www..... 60
OY 1604 tacaatatataattagatactagtatgyatattccaataacataacttggalgttta 16633
Db 61 www..... 120
OY 1664 aactaatcttgtttcctcctaagcataaataatcatcatcgaggtaaaaaagtttg 17233
Db 121 www..... 180
OY 1724 tcttatcttcgcatgaagaagaaacctaatgaacttaatttttgaanaagtaaac 17833
Db 181 www..... 240
OY 1784 cctttaccatagatlaattaccogtatgttttggtcgaataagcacctctacaac 18431
Db 241 www..... 300
OY 1844 tgyatagtcgaatttttctgcaaalataaattaggaaatcoaatgylactatacaga 19033
Db 301 www..... 360
OY 1904 agaaacgcctgagatataacctttaacttaataagaacaaaattttgaaaaagttataact 19633

```

Db	361	www.....gctta.....	420
Oy	1964	tctacaatatataaataatgacctataatgattctctaagtcttaaatatt	2023
Db	421	www.....	480
Oy	2024	tttttattattgtaataacatgtaaaccaataagtgtgtaattcaaatc	2083
Db	481	www.....	540
Oy	2084	tccataataatttggaaatctacaatbataataattgtagcaataaatagcataa	2143
Db	541	www.....	600
Oy	2144	aagttccaaaaaaatttgttcaacgaactccaaattttttttttagagaaacag	2203
Db	601	www.....	660
Oy	2204	aataacagatagaanaactatttgttggaaatgtagtaatatatacatgaacaa	2263
Db	661	www.....	720
Oy	2264	ttttaaaaaatataagacctaaagcgctcaagatagtatctagttagtgaatta	2323
Db	721	www.....	780
Oy	2324	ataat 2328	
Db	781	www 785	

RESULT	9
ID	AAF58254
AC	AAF58254 standard; DNA; 936 BP.
XX	
AC	AAF58254;
XX	
DT	24-APR-2001 (first entry)
XX	
DE	Oligonucleotide D1875.
XX	
KW	Electron-transfer group; ETM; mismatch; genotyping;
KW	gene expression; ss.
XX	
OS	Synthetic.
XX	
PN	WO200107665-A2.
XX	
PD	01-FEB-2001.
XX	
PF	26-JUL-2000; 2000WO-US20476.
XX	
PR	26-JUL-1999; 99US-0145695.
PR	17-MAR-2000; 2000US-0190259.
XX	
PA	(CLIN-) CLINICAL MICRO SENSORS INC.
XX	
PI	Unex RM;
XX	
DR	WPI: 2001-159728/16.
XX	
PT	Nucleic acids containing electron-transfer group, useful as labels in
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT	a single surface -
XX	
XX	
PS	Example 6; Page 127; 159pp; English.
XX	
CC	The present invention relates to a composition comprising two nucleic
CC	acids each containing an electron-transfer group (ETM) having
CC	different redox potentials. The invention is used for electronic
CC	detection of nucleic acids, especially of substitutions (mismatches)
CC	and single-nucleotide polymorphisms, e.g. for genotyping,
CC	monitoring gene expression.

XX Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;
SQ

Query Match	3.6%	Score 250.6	-DB 22	Length 936
Best Local Similarity	0.5%	Pred. No. 1.4e-30		
Matches 4	Conservative 396	Mismatches 185	Indels 0	Gaps 0

QY	1544	taagtcctttaacattatgttgaattgaattaaagcgaaactgtaaacat	1603
Db	1	#####	60
QY	1504	tacaatattatattagatactagtagtgatattccaatacacttgtagttta	1663
Db	61	#####	120
QY	1664	aactaatctgtttctcctaagctataataatatacactgaaggtaaaaagtttg	1723
Db	121	#####	180
QY	1724	tcttatttcgcatgcatgaagataaaccctaactgaacttaatttttggaaatgtaac	1783
Db	181	#####	240
QY	1784	cctttactcatagatgaattacactgtagttttgttgcataatagacagcctctacac	1843
Db	241	#####	300
QY	1844	tgtgatgtcaatttttctgcgaataataaattaggaatcaatgctactatactga	1903
Db	301	#####	360
QY	1904	agaaacagctgagatatacatttaattaaagacaaattttgaaaatgtaatt	1963
Db	361	#####	420
QY	1964	tctacaatatattaataatagatgctataatgatttcctatgttcttaaaattt	2023
Db	421	#####	480
QY	2024	tttttattatgtctataatacatataigaaccaataatgttggtgaattcaaatc	2083
Db	481	#####	540
QY	2084	tccattaatatttttgaatacctcaaaattatattatattagtcataaagatgcataga	2143
Db	541	#####	600
QY	2144	aagttccaaaaaaatttgttaacagaaactccaatttttttttaltgagacaag	2203
Db	601	#####	660
QY	2204	aaataacagatagaaacttttgttgygaatggaagttagtaatacatataagcaaa	2263
Db	661	#####	720
QY	2264	ttttaaaaaattatataagcctacagcgtcaagatggtatctagtaggtgtaatta	2323
Db	721	#####	780
QY	2324	ataat 2328	
Db	781	#####	
RESULT	10		
AAFS8257			
ID	AAFS8257	standard; DNA; 936 BP.	
XX	AAFS8257:		
XX			
XX	24-APR-2001	(first entry)	
XX			

DE Oligonucleotide D1954.

KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.

OS Synthetic.

PN WO200107665-A2.

PD 01-FEB-2001.

PF 26-JUL-2000; 2000WO-US20476.

PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0100350

XX
DA /CIT FN-) CIT TNGAT MTQBO CENCO

XX
XX
BT
Time: PM:

XX
DP: 30

XX	Nucliole acide
PT	consta

PT hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface.

```
XX      Example 6: Data 127: 15000: Esc14cb
PC
```

CC The present invention relates to a composition comprising two nucleotides
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.

50 Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

Query Match	3.6%	Score 250.6	-DB 22	Length 936
Best Local	Similarity 0.5%	Pred. No. 1.4e-30		
Matches	4	Conservative 556	Mismatches 185	Indels 0
				Gaps 0

Oy	1544	taagcaatttaacattatgttgaatgtgaatttaagaacgactgtacaat	1603
	1	60
Oy	1604	tacaataattatataagatactagatgtaattatccaaatacaacttggatgtta	1663
	61	120
Oy	1664	aacttaactgtttctctcttaagatataatataatcatcagaggttaaaagtttg	1722
	121	180
Oy	1724	tctattttcgagatgaatgaagataaacctaacttaatttttggaaaatgtaac	1783
	181	240
Oy	1784	ccatttactatagataatataacgcatgtttttgtgcataaagacagcctctacaac	1843
	241	300
Oy	1844	tgtatagtcgaatttttcgcgaatatbaatttaggaatcgaatgcatactacataga	1903
	301	360
Oy	1904	agaacagcttgatatacatatttaatttaagaacaaaattttgaaaaatgltataat	1963
	361	420
Oy	1964	tctacaatatatataaaataatgatgcctataatgtaatttctctatgtttcttaaatatt	2023
	421	480

Oy	2024	tttttaaattagtttaataatcattgaaccaataaagtgttgtagatccaatc	2083
Dd	461	www.....	540
Oy	2084	lccattaatatttttgaactctacaatatbaattatattagtcataacaaatgcataga	2143
Dd	541	www.....	600
Oy	2144	aagttccaaaaaaattttgttaacgaaactcccaatttttttttatatggaacaag	2203
Dd	601	www.....	660
Oy	2204	aataacagatagaaaaactatttgtttgtgaagtgaagtagtaatacatgaacaa	2263
Dd	661	www.....	720
Oy	2264	ttttaaaaaatlatagaacctatacgctcacaaagtatgtatctagtgaagtgaatta	2323
Dd	721	www.....	780
Oy	2324	ataat 2328	
Dd	781	www 785	
RESULT 11			
ID	AAFS8259	standard; DNA; 936 BP.	
XX	AAFS8259;		
XX	AC		
XX	DT	24-APR-2001 (first entry)	
XX	DE	Oligonucleotide D2004.	
XX	RW	Electron-transfer group; ETM; mismatch; genotyping; gene expression; ss.	
XX	OS	Synthetic.	
PN	WO200107665-A2.		
PD	01-FEB-2001.		
PF	26-JUL-2000; 2000MO-US20476.		
PR	26-JUL-1999; 99US-0145695.		
PR	17-MAR-2000; 2000US-0190259.		
PA	(CLIN-) CLINICAL MICRO SENSORS INC.		
PI	Umek RM;		
Pt	WP1: 2001-159728/16.		
Pt	Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface		
PS	Example 6; Page 128; 159pp; English.		
CC	The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.		
SO	Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;		
Query Match 3.6%; Score 250.6; DB 22; Length 936;			

```

Best Local Similarity 0.5%; Pred No. 1.4e-30;
Matches 4; conservative 596; Mismatches 185; Indels 0; Gaps 0;

QY 1544 taatgacatttaacaatattatgttctggaattcaataagsgaactgttaacat 1603
    :::: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 60

QY 1604 taacaatttatatgaactagatgttgatctccaataacatacttgatgttta 1663
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 120

QY 1664 aacttaactgttcttctcctacggtataaataatcatcagagtaaaaaagtgtg 1723
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 180

QY 1724 tcttatcttcgcgatgcatagaagsgataaacctaagtacttaatttttgaaaatgtaac 1783
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 240

QY 1784 cctttactacatagataaatacgaatgttttctgtgcabaatgacagcctcacaac 1843
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 300

QY 1844 tgtgtagtagcaatttctctcgcaaatataatagaatcaatgcatactacaataga 1903
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 360

QY 1904 agaaacagctgtagtattacatttaattaaagacaaaattttgaaaaatgttaatt 1963
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 420

QY 1964 tctaacaaatattataaataatgatgactataatgtattcctatgtcttaaaatatt 2023
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 480

QY 2024 tttttataattagattaaataacattatgaacaaataatgttggtaattcaaatatc 2083
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 481 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 540

QY 2084 tccattaatatttttgaatcacaataatataatttagtcaataacaatgataga 2143
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 541 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 600

QY 2144 aagttccaaaaaaatttgttaacagaactccaatttttttttatggaacaag 2203
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 601 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 660

QY 2204 aataacagatagaacaaactattgttctggaatgaagtagtaatacatatgaacaa 2263
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 661 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 720

QY 2264 ttttaaaaaatataatagcctatacgcgctcaaaagtatgtatcagtgagtgtaatta 2323
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 721 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 780

QY 2324 ataatt 2328
    : : : :
Db 781 www.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww.ww 785

RESULT 12
AAFS8262
ID AAFS8262 standard; DNA; 936 BP.
xx
AC AAFS8262;
xx
xx 24-APR-2001 (first entry)
DT
xx Oligonucleotide D2007.
DE
xx
xx Electron-transfer group; ETM; mismatch; genotyping;
KM gene expression; ss.
xx

```

OS Synthetic.
 PN WO200107665-A2.
 PD 01-FEB-2001.
 XX 26-JUL-2000; 2000WO-US20476.
 PF 26-JUL-1999; 99US-0145695.
 PR 17-MAR-2000; 2000US-0190259.
 XX
 PA (CLIN-) CLINICAL MICRO SENSORS INC.
 XX
 PI Umek RM;
 XX
 DR WPI; 2001-159728/16.
 XX
 XX Nucleic acids containing electron-transfer group, useful as labels in
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
 PT a single surface -
 XX
 PS Example 6; Page 128; 159pp; English.
 XX
 XX The present invention relates to a composition comprising two nucleic
 CC acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 CC detection of nucleic acids, especially of substitutions (mismatches)
 CC and single-nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.
 XX
 XX Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other;

[illegible]

```

RESULT 13
AAF58255 standard; DNA; 938 BP.
ID AAF58255
XX AAF58255;
AC AAF58255;
XX 24-APR-2001 (first entry)
DT 24-APR-2001 (first entry)
XX Oligonucleotide D1876.
DE Oligonucleotide D1876.
XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX Synthetic.
OS Synthetic.
XX WO200107665-A2.
PN WO200107665-A2.
XX 01-FEB-2001.
PD 01-FEB-2001.
XX 26-JUL-2000; 2000WO-US20476.
PF 26-JUL-2000; 2000WO-US20476.
XX 26-JUL-1999; 99US-0145695.
PR 26-JUL-1999; 99US-0145695.
XX 17-MAR-2000; 2000US-0190259.
PR 17-MAR-2000; 2000US-0190259.
XX (CLIN-) CLINICAL MICRO SENSORS INC.
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX Umek RM;
PI Umek RM;
XX WPT: 2001-159728/16.
DR WPT: 2001-159728/16.
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PP a single surface -
XX a single surface
XX Example 6; Page 127; 159pp; English.
PS Example 6; Page 127; 159pp; English.
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
CX monitoring gene expression.
XX monitoring gene expression.
SQ Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

Query Match          3.6%; Score 250.6; DB 22; Length 938;
Best Local Similarity 0.5%; Pred.No. 1.4e-30;
Matches      4; Conservative 596; Mismatches 185; Indels    0; Gaps     0;
OY   1544 taatgcattcaacatatcttactgttagtggaatttaataagaagcgaaacttgtaacat 1603
       :::: :::::::::::: :::: :::::::::::::: ::::: ::::: ::::: :::::

```

```

Db 1 www..... 60
Qy 1604 tacaataattatagatactagatgtatattccaaatacatttgatgtta 1663
Db 61 ..... 120
Qy 1664 aacttaactgttcttccctacgataaataatcaacgcaggtaaaaagtttg 1723
Db 121 ..... 180
Qy 1724 tcttatttcgagatgaagataaacctaagcttaattttttgaaaaatgtaac 1783
Db 181 ..... 240
Qy 1784 ccttactacatagataatcagatgttttgtgtccataatgacagcctacac 1843
Db 241 ..... 300
Qy 1844 tctgatagtcaattttctcgcataataatagaatcaatgctactcaataga 1903
Db 301 ..... 360
Qy 1904 agaacagctgagatfacatttaataagaacaaattttgaaaaatgtataatt 1963
Db 361 ..... 420
Qy 1964 tctaacaataatataaataatgctcctaatagtattcctcctaaataatt 2023
Db 421 ..... 480
Qy 2024 tttttcatttgcataataacatatagaaccaataaagttgtgaattcaaatc 2083
Db 481 ..... 540
Qy 2084 tccattcaatttttgaatactcaaaatttaatttaagtcataaacaatgcataga 2143
Db 541 ..... 600
Qy 2144 aagttccaaaaaatttgttaacagaaactccaaatttttttttatagaacaag 2203
Db 601 ..... 660
Qy 2204 aaataacagatagaaaactatttgtgtggaatggaagagtaataacataaagcaaa 2263
Db 661 ..... 720
Qy 2264 tttaaaaaatataaagcctacgcgtcaagaatgtatctagtagtgaatta 2323
Db 721 ..... 780
Qy 2324 ataact 2328
Db 781 ..... 785

RESULT 14
ID AAT85307 standard; DNA: 4937 BP.
AC AAT85307;
XX
XX
DT 14-APR-1998 (first entry)
XX
XX Arabidopsis thaliana cytochrome P450-type hydroxylase DNA.
DE
XX
XX Cytochrome P450-type hydroxylase; identification; brassinosteroid;
KM brassinosteroid inhibitor; modified plant; recombinant production;
XX teaesterone; ds.
OS Arabidopsis thaliana.
XX
XX Key Location/Qualifiers
FH CDS 966..4773
FT

```

```

FT FT /*tag= a
FT FT /product= cytochrome_P450-type_hydroxylase
FT FT 1484..1679
FT FT /*tag= b
FT FT /number= 1
FT FT 1830..1916
FT FT /*tag= c
FT FT /number= 2
FT FT 2166..3902
FT FT /*tag= d
FT FT /number= 3
FT FT 3990..4083
FT FT /*tag= e
FT FT /number= 4
FT FT 4163..4247
FT FT /*tag= f
FT FT /number= 5
FT FT 4355..4445
FT FT /*tag= g
FT FT /number= 6
FT FT 4577..4673
FT FT /*tag= h
FT FT /number= 7
XX PN MO9735986-A1.
XX PD 02-OCT-1997.
XX XX
XX PF 27-MAR-1997; 97MO-EP01586.
XX PR 27-MAR-1996; 96US-0622166.
XX XX
XX PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX XX
XX PI Altmann T, Koncz C, Mathur J, Szekeres MA;
XX XX
XX DR WPI; 1997-489649/45.
XX DR P-PSDB; AAM27153.
XX XX
XX PT New isolated plant cytochrome P450-type hydroxylase gene - used to
XX PT identify substances acting as brassino-steroid(s) or brassinosteroid
XX PT inhibitors for the production of modified plants
XX PS Claim 1; Pages 48-53; 77pp; English.
XX XX
XX CC The present sequence encodes Arabidopsis thaliana cytochrome
XX CC P450-type hydroxylase. The hydroxylase can be used to identify
XX CC brassinosteroids or brassinosteroid inhibitors, useful to produce
XX CC plants with modified physiological and/or phenotypic
XX CC characteristics. The modified plants may show, e.g. stimulation of
XX CC growth, increased cell elongation, increased wood production,
XX CC accelerated seed germination at low temperatures, an increase in
XX CC dry weight, repressed anthocyanin production during growth in light
XX CC and/or inhibited de-etiolation which is induced, e.g. by cytokinin,
XX CC in the dark or an increase in stress tolerance. The hydroxylase or
XX CC its coding sequence can also be used for the recombinant production
XX CC of compounds, e.g. teaesterone.
XX SQ Sequence 4937 BP; 1526 A; 864 C; 870 G; 1677 T; 0 other;

Query Match 1.3%; Score 92.8; DB 18; Length 4937;
Best Local Similarity 53.1%; Pred. No. 4.1e-06;
Matches 225; Conservative 0; Mismatches 192; Indels 7; Gaps 1;
Qy 3501 caggtatggaagaatataatagatcgactgtttggagaacacagatcgtatcagctga 3560
Db 1162 ccggtacggtcgtgttccatgacgactcttttggtagcagcagatttccagctga 1221
Qy 3561 tctgacttaataatgatacattacaacgaaggaagcctcttgaaatgtagtatcc 3620
Db 1222 ccggaacgaacgcggtttgttcttcgaagcgaagggaagcgttttggagtggtcttatcc 1281

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2001, 06:49:51 ; Search time 8875.91 seconds
(without alignments)
12003.485 Million cell updates/sec

Title: US-09-502-426-1
Perfect score: 6888
Sequence: 1 atgtgggtattatattgttg.....cctagtcagcgccaagctt 6888

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_on:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_on:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v1:*
59: gb_v12:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rot1:*
95: gb_rot2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
C 1	6753.6	98.0	84196	14 ATT3A5	AL132979 Arabidops
2	4818	69.9	4818	12 AF044216	AF044216 Arabidops
3	134	1.9	100269	13 AF18022	AF18022 Arabidops
4	133.2	1.9	104992	60 AC005504	AC005504 Plasmodiu
5	133.2	1.9	169546	60 AC004157	AC004157 Plasmodiu
C 6	130.8	1.9	104992	60 AC005504	AC005504 Plasmodiu
7	130.8	1.9	169546	60 AC004157	AC004157 Plasmodiu
8	129.4	1.9	14867	4 AE001398	AE001398 Plasmodiu

9	128	1.9	106142	12	AC025417	AC025417 Genomic s
10	127.8	1.9	86827	96	PFMAL3P5	AL034556 Plasmodiu
11	125.4	1.8	86827	96	PFMAL3P5	AL034556 Plasmodiu
12	124.4	1.8	4601	6	DMU1584	U1584 Drosophila
13	124.4	1.8	19517	6	DMU37541	U37541 Drosophila
14	123.4	1.8	67970	96	PFMAL1P3	AL031746 Plasmodiu
15	122.6	1.8	4601	6	DMU11584	U11584 Drosophila
16	122.6	1.8	19517	6	DMU37541	U37541 Drosophila
17	118	1.7	67970	96	PFMAL1P3	AL031746 Plasmodiu
18	118	1.7	84196	14	ATT3A5	AL132979 Arabidops
19	113	1.6	14867	92	HS1108D11	AE001398 Plasmodiu
20	107.8	1.6	152209	92	HS1108D11	AL034419 Human DNA
21	105.6	1.5	176174	86	AC007483	AC007483 Homo sapi
22	103	1.5	130117	85	AC004907	AC004907 Homo sapi
23	103	1.5	149733	77	AC087568	AC087568 Pan trogl
24	102.6	1.5	15643	5	AF149768	AF149768 Bombyx mo
25	101.4	1.5	178087	85	AC005089	AC005089 Homo sapi
26	101.2	1.5	15421	96	PRCOMPTA	X95275 P.falciparu
27	100.8	1.5	152209	92	HS1108D11	AL034419 Human DNA
28	100.8	1.5	188450	79	AL158209	AL158209 Homo sapi
29	99.6	1.4	180741	75	AC073846	AC073846 Homo sapi
30	99.4	1.4	1867	14	MTCARJ23	AJ223323 Saccharom
31	99.4	1.4	175223	85	AC004617	AC004617 Homo sapi
32	99.4	1.4	310779	60	AC005140	AC005140 Plasmodiu
33	99.9	1.4	185603	74	AC073136	AC073136 Homo sapi
34	98.4	1.4	2290	6	MIDYTRN	X05915 D. yakuba m
35	98.4	1.4	16019	6	MIDYTRN	X03240 Drosophila
36	98.2	1.4	12029	4	AE001427	AE001427 Plasmodiu
37	97.4	1.4	62268	92	HS1178121	AL109852 Human DNA
38	97.2	1.4	159781	69	AC025018	AC025018 Homo sapi
39	97	1.4	159593	85	AC004832	AC004832 Homo sapi
40	96.8	1.4	159599	75	AC074354	AC074354 Oryza sat
41	96.6	1.4	234112	96	PFMAL4P2	AL035475 Plasmodiu
42	96.4	1.4	242893	83	CEY53C12	Z92859 Caenorhabet
43	96	1.4	146285	85	AC005083	AC005083 Homo sapi
44	95.8	1.4	155012	78	AL133476	AL133476 Homo sapi
45	95.6	1.4	156060	60	AC004153	AC004153 Plasmodiu

ALIGNMENTS

RESULT 1	ATT3A5	84196 bp	DNA	26-JAN-2000
LOCUS	Arabidopsis thaliana	DNA chromosome 3,	BAC clone T3A5.	
DEFINITION	AL132979			
ACCESSION	AL132979.2	GI:6782244		
VERSION				
KEYWORDS	thale cress.			
SOURCE	Arabidopsis thaliana			
ORGANISM	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE	1 (bases 1 to 84196)			
AUTHORS	Bloeker, H., Mewes, H.W., Lemcke, K., Mayer, K.F.X., Quetier, F. and Salanoubat, M.			
JOURNAL	Unpublished.			
REFERENCE	2 (bases 1 to 84196)			
AUTHORS	EU Arabidopsis sequencing, project.			
TITLE	Direct Submission			
JOURNAL	Submitted (25-JAN-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project Coordinator: Marcel Salanoubat and Francis Quetier, Groupement d'Interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue Gaston Cremieux, BP191, 91006 Evry Cedex, France; http://www.genoscope.cns.fr			
COMMENT	On Jan 27, 2000 this sequence version replaced gi:6434247. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/. Location/Qualifiers			
FEATURES				

source	1. .84196	/organism="Arabidopsis thaliana"
		/variety="Columbia"
		/db_xref="taxon:3702"
LTR	2983. .3338	/chromosome="3"
misc_feature	/note="355 bp LTR"	
LTR	3339. .7564	/note="LTR-transposon"
gene	7565. .7922	/note="367bp LTR"
CDS	12339. .13197	/gene="T3A5.10"
	join(12339. .12665,12742. .12803,12888. .13055,13125. .13197)	/gene="T3A5.10"
	/note="similarity to cyclin-dependent kinase inhibitor (ICK1), Arabidopsis thaliana, EMBL:AT094772.1"	
exon	/codon_start=1	/product="putative protein"
intron	/protein_id="CAB62432.1"	/db_xref="GI:6561966"
exon	12742. .12803	/number=1
intron	/gene="T3A5.10"	/number=2
exon	12804. .12887	/gene="T3A5.10"
intron	/number=2	
exon	12888. .13055	/number=3
intron	13056. .13124	/gene="T3A5.10"
exon	13125. .13197	/number=3
intron	/gene="T3A5.10"	/number=4
exon	15858. .16470	/number=1
gene	/gene="T3A5.20"	/complement(15858. .15901)
CDS	/note="similarity to predicted proteins, Arabidopsis thaliana"	
	/codon_start=1	/product="putative protein"
	/protein_id="CAB62433.1"	/db_xref="GI:6561967"
	/translation="MARLEHNSMNGAARISFSENFVEIRSEKSNKNNINSRSSFS MSADAFSVTDVSMIPADEIFLFGKILPFKETSIVRTLGEEILTEEGSMWDGTF SRPILSSSFSTKGTWRELLGLKRTVRSKTKTDVNEVGLSQDKIISGNVATREC QVADPR"	
intron	complement(15902. .16013)	
exon	/number=1	
gene	complement(16014. .16470)	
exon	/number=2	
exon	18303. .19931	/gene="T3A5.30"
exon	complement(18303. .19931)	
CDS	complement(18303. .19931)	
	/codon_start=1	/product="scarce row-like 7 (SCL7)"
	/protein_id="CAB62434.1"	

/db_xref="GI:6561968"
/translation="MAYMTDSCNLMIAAOOLIKOQOQSOHQOQEEQOEQENPMFN
PSRGTLPSSGSPFQVNTDPCGHFPHLEHNOHAAVASEEPDSEI.INGDS
QTPDPPIVGHDPVSPSPRLSAPSYLNRYNKDSSAOCAPPASATIAISSPSRQ
HPPPPPPDPDLPNPIFEKAIHDYARKPEKPDLLIKESVSESDPTORVGTFAE
ALSHKEESSSSSSSSLEDFILSYKTLNACPSKPAHLTANOAILLEANOSNNIHI
VDGIFQGIOMSLALQALATRSSGKPRIRISGIPAPSLGSDSPSLIATGNRLRDA
AIDLNEEFYPLVPIQLNGSSFRVDEVLYVFMLELYKLDEATATVYALRA
LSNPRIVLTGEYVSLNRYEPANRYKNSLRFSAPFESLEPNIDRSKRLRYEVL
EGRIMDLVSSDDNNKPTGRCFIMEKEQRMVLEKAGEFEPKPSNVAVSQAKLLIM
NYSSTLSLVSESPGFI SLAMNVPILTVSSMR"
/gene="T3A5.40"
/complement(join(26093..26338,27091..27200,27308..27386,
27478..27570,27723..28037,28136..28288,28374..28658,
28779..28999))
/note="Contains Cytochrome P450 cysteine heme-iron ligand
signature AA455-464; Prokaryotic membrane lipoprotein lipid
attachment site AA320-330"
/codon_start=1
/product="steroid 22-alpha-hydroxylase (DMF4)"
/protein_id="CAB62435.1"
/db_xref="GI:6561969"
/translation="MEETEHHTLPLLLPSLSLLPLILKRNKRTFENLPGRKS
GMPFLGETIYLKPYATTLGDEMOCHSVKGYKLYRSNLGEPITVADAGLNFILO
NEGRLFECSYPSRISGILGKWSMLVGDHMRDRSISLNPISARLRTILKQVEH
TLFLYDSMOQNSIFGSADDAKKEPTENMAKHIMSDPEEBTBOUKREYITPMKGVY
APLNPOTAYHKALOSRAITLKTIERMEERKLDIKEDDEEVEKDEENKSDH
VKRQRTDDLLGWVLKHSNLSLEQILDILSLFAGHETSVAIALAIFPLQAPKAV
EELREHEILIRAKKEGSESLMDYDKKDFQCVINETLRGNVRFHRAKIAKY
RYKGYDIPSGMKLPVISAVIDNSRDENLFPNMOQONGASSSGSFETGWN
NYMFGGGRPLCAGSELAKLEMAVEIHLVLKFWELAEDDKPAFEPVDFPNLPIR
VSRLL"
exon
/complement(26093..26338)
intron
/complement(26339..27090)
/number=1
exon
/complement(27091..27200)
/number=2
intron
/complement(27201..27307)
/number=2
exon
/complement(27308..27386)
/number=3
intron
/complement(27387..27477)
/number=3
exon
/complement(27478..27570)
/number=4
intron
/complement(27571..27722)
/number=4
exon
/complement(27723..28037)
/number=5
intron
/complement(28038..28135)
/number=5
exon
/complement(28136..28288)
/number=6
intron
/complement(28289..28373)
/number=6
exon
/complement(28374..28698)
/number=7
intron
/complement(28699..28778)
/number=7
exon
/complement(28779..28999)
/number=8
gene
38307..41323
/gene="T3A5.50"
/complement(38307..38796)
exon
/number=1
CDS
/complement(join(38307..38796,38877..39090,39189..39270,
40180..40242,40337..40401,40489..40564,40655..40734,
40805..40919,41325..41323))
/note="Contains Eukaryotic putative RNA-binding region
RNP-1 signature AAI/9-186"
/codon_start=1

/product="U1 snRNP 70K protein"
/protein_id="CAB62436.1"
/db_xref="GI:6561970"
/translation="MDSGPPFLRNPAVQARAKVONRANVQLKMGSHPLGLTN
NLKLFEPREPLEKPPPERKPCPTGMAQFVSNFEPDDEPAEPKPEVLEPSQR
ERHILKLEGEVEKAAEDLKYPDNPNPATGDPKTLFVSLRYESESSEKIKKEFS
YGPTRVHLVTDQTLNPKGVALEFYHNTDMKAAVQAQOGKIDGRVIVYERGT
VPMNRPRLRGSGSTSRVGGSEVGGQOGRTSSESPRSREERKREKERE
RSRLSHEQREERSRDRPRDDKHHDRDGGRRDRSDRDRDRDRDRDRDRDR
GRDRTSHDRDRDRSRKREKERYEGSEYHEGGGGRSREDAERYRGEPTKGYEDDQGD
TDKSHRYDYDEEDDFRYEREYKRKRSSEREYVR"
/gene="T3A5.40"
/complement(38877..39090)
/number=1
exon
/complement(39091..39188)
/number=2
intron
/complement(39189..39270)
/number=3
exon
/complement(39271..40179)
/number=3
intron
/complement(40180..40242)
/number=4
exon
/complement(40243..40336)
/number=4
intron

Query Match 98.0%; Score 6753.6; DB 14; Length 84196;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 6864; Conservative 0; Mismatches 19; Indels 9; Gaps 9;

1 atgtggatcatatattgttgggttcggtttgagccacataataattcgtttcgtt 60
Db 32199 ATGTGGATCATATATTGTTGGGTTGCGTTGAGCCACATATAATTTCGTTCTGCTGT 32140

61 tatctgtccatgatattgagttggttcctcaatttgatccagaataataat 120
Db 32139 TATCTGTTCACATGATTTGAGTTGTTGTTCAATTGATTCGAAGATAATTAATATTT 32080

121 aaatccttaataatattcaagtaataatattatcttcatattgttttataa 180
Db 32079 AAAATTCATTTAAATTAATTTCAAGTAATTAATTTGTTACATTTGTTGTTATTAACA 32020

181 aatcctatcttggatattagaaataatgagttggaatttaataataaagaa 240
Db 32019 AATATCTATCTTTGATATATGAGAAATATGAGATTGATTAATTAATTAAGGAAA 31960

241 taatcgtatccattggttgattacacagttaagtttctgttcttctgtata 300
Db 31959 TAATCGATTCCATTTGTTGATTAACAGTTAACTTTTGTGTTCTTTGTTATATGT 31900

301 atatgagtaaatcaaaagagatattgaaagtgtaaacatatttggttaag 360
Db 31899 ATATGAGTAAATCAAAAGACATATTGAAAGTAAACATATTTCGTTATGACCCCC 31840

361 aaaaaaataaaaaaacaacaacaacccccccgatatagtttgggttcgga 420
Db 31839 CAAAAAATAAAAAAACAACAACAACA -ACAATTTTCATATAGTTTTCGTTTCGGAAT 31781

421 aggttatttgatcaataatcatgcatcattcttcttgattactaagaatt 480
Db 31780 AGGTTATTGATCATTAATTCATGATCATTCATTCTTATTCATTAAGAATTTCTTAC 31721

481 caatcaaaatttgatcatatctctgattataataataacggttggaatcc 540
Db 31720 CAATTAATAATTCGAATTCATTCCTTGATTAATTAATTAATGAAGTGAATATCC 31661

541 gttatcgatcacatcaatcatgatattctgtgtctaaatccagaatattaa 600
Db 31660 GTTATTCGATCATCTCCATCATGATTATGATCTTGCTGCTAATCCAGAAATATTATACA 31601

601 agagtatcagagaaaaacccaataaagaagaagagtagtgaccatgagat 660
Db 31600 AGAGTATTGAGAAAAAAGCAAAATTAAGAAAGAGAGAGACCCCATGAGATG 31541

QY	661	tgaaataatcatcaaagagaa	taagagatgacaa	ccaaagggtgtggaataatggtccct	720
Db	31540	TCGAATATTATATCAAAAGAGATAAGAGATGACAAACCAAAAGGTTGGTAATATGATGTCCT			31481
QY	721	ggcaagcttctctccaaatcaaat	catccaccatcttgga	ttctctcgtatctcgttaaa	780
Db	31480	GCCAGCTTCTCTCAACAAATATCGAACCCATTGTGGATTTTCTGGAATATTCGTTMAAA			31421
QY	781	tttgagataaagatctgtgga	aaaaatcttatttgtta	gctgaatctcaatattatgttcaa	840
Db	31420	TTTGGCATTAACCATGTGTGAAAAATTTTTATTGTAGCTGTATCAATATATATGTTTCCA			31361
QY	841	ggatattgataaatctctcgtt	taagaaatcttgccctctctt	gttctgttccctccct	900
Db	31360	GGTATTTGCAATATCTCTGTTTAAAGCATATTTGTCTTTTGTGTTTGATTTCTCT			31301
QY	901	taactatataatctcgcgga	tatatgatacaatgatatata	catcaaaacaatctgtctgg	960
Db	31300	TAACTATATATATATCGCGATATATGATTAACATGATATATATCAAAACAAATATGATCTGGG			31241
QY	961	accattttgaaataaactttt	ctcaacatctacgggaacat	ctggaatcgacccttaata	1020
Db	31240	ACCATTTTGAATTAACCTTTTCTCAAACTTACGGGACATCTGACTCGCAACCTT-AAATTA			31182
QY	1021	cgatttcaagcgctcaatctg	tgaatctgaatctgaataagca	taaggaacccgttcaagc	1080
Db	31181	CGATTTTACAGCGTCACTAGTGTGACATTTCTAGCATTAATTAAGCATTAAGGACCCGTTTCAAGC			31122
QY	1081	tatttatacaagcttcaaaa	atcgaatagctttgaaatcc	ctttagaaaaatttggaatba	1140
Db	31121	TATTTTATCAAAAGTTTACAAACGTAATATATGCTGTGAAATCTTTTACAAAAATTTTGGAAATTA			31062
QY	1141	ccgattgttcatgtaaataga	tttagtctgtgaatacaaatatg	tttaactaattagtggtca	1200
Db	31061	CCGATTGTATTGTAAATATATAGATTATAGTGTGTAAACAAATATTTTATCATCAATTAGTGTGCA			31002
QY	1201	acatacacataatcccttca	cagaaaaaacaactaagagaag	cttcaacatcatcatat	1260
Db	31001	ACATATATACATTAATTCCTTACAGAAAAAACAACTTAAGGAGATTACATATATCCATATAT			30942
QY	1261	gggtatgcatatccctttctc	gafatgctatactacgaagac	taagaatagtatgtagtctc	1320
Db	30941	GGGTATATGCTATACCTTTCACGTATGCTATATATAGAGACTTAAAGTAATATGATATATGTC			30882
QY	1321	gataaatacgaatactcaac	gcgcgtgtcaataatata	tggaacgcatgttcaacatcatgcaa	1380
Db	30881	GATTAATATGAATTTCAACGCCGTGTATATATATGTGGACCGTATGTTCAGATCACTGCAG			30822
QY	1381	atacaatctctgtgtgtgt	ccaacataaaaaacaaga	aaaaaagaagaacgatcttt	1440
Db	30821	ATATCATTTCTTGTTGGTGTCTACAAATTAATAAACAACAAAAACAAGAAAAACGATTTTTT			30762
QY	1441	cttggaatccatctcaatagt	ctcaaaatgataatctt	tttggtaagctttcgaaagtc	1500
Db	30761	CTTGGAATTCATATCAATGTATCTTAATAATGATATGATCTTTTGGGTTTACAGTTTTCGAAGTTC			30702
QY	1501	ctcaacaagcgtttaacca	ctctgcaacatctaaat	tgctctctctcttcaatcatcttaacat	1560
Db	30701	CTCAACAAGCGTAAACCAATCTGCAACTATTAATTTGCTTTTATATGATGATTTTAACAT			30642
QY	1561	atttatctgttaagtgtgga	attcaataagcgaactct	gttaacatctcaatatattataatag	1620
Db	30641	ATTATATGTTTAACTTGGAAATTTAATAAGACGCAACTGTGAACATTTCAATATATTATATTAG			30582
QY	1621	atactatgatactgataat	ctccaatacaatcttgga	tgatttcaacttaaatctgttctc	1680
Db	30581	ATACATATGATATGATTTATTCCAAATATATCAATCTTGTGAGTGTTTAAACTTATATCTGTTTCT			30522
QY	1681	ccctacggtataataata	taacatcgaaggttaaaaaa	agtttgttccttaatttcgcgaatgc	1740
Db	30521	TCTCTAC-GTATTAATAATTTATATATGATGAGGTATAAAAAAGTTTGTCTTATATTTTCCGATATGC			30463

QY	1741	atgaagagataaacttaatgacttaatttttttggaaaatgtaacccctttactcagaagt	1800
Db	30462	ATGAAGGATTAACCTAATGACCTTAATATTTTGGAAATATPAACTCCCTTTACTCATAGATT	30404
QY	1801	aattacgcatgcttcttcttggccaataagacagccctcctaactctgtagtcaattctt	1860
Db	30402	AATTACCGTATGTTTTGTTGGCATATATGACACGCTCTCAACGTGTGATGACAAATTTTT	3034
QY	1861	tcgcaaatatlaaataggaattcaatgctactactcaatagaagaagaacagctgagatc	1920
Db	30342	TCTGCAAATATTAATTAATGGAATTCATAGCTACTATCAATTAACAAACAGCTGTGATTT	3028
QY	1921	acattttaatlaaagacaaactttctgaaaaatgctaaactcttcaacatatacttaa	1980
Db	30282	ACATTTAATTAAAGCAAAATTTTTGAAAAATGTTAATATTCTCAACANATATATTAA	30222
QY	1961	aataatgectlaaataatgatacttccctaagcttcttaaatattttcttaataatgta	2040
Db	30222	AATATGATGCTTAATTAATTAATGATTTCCATATGTTCTTAATAATTTTTTTTAAATATTAGTAA	3016
QY	2041	taaatacatlaatgaacaaataatagtttgggaatcaaatatctcacttaattatttgg	2100
Db	30162	TAAATACATTATGACCAATTAATATGTTGGTAATCAAAATATCTCATTAATATTTTTTG	3010
QY	2101	aaatccaaactlaataatlttagtcaataaacaatgcaatagaagcttccaaaaaaat	2160
Db	30102	AAATCTACAAATTTATTAATTTATTAAGCAATTAACAAATGATGAAGTTCCAAAAAATTT	3004
QY	2161	ttgttaacagaacttccaaatttttttttttttttttttttttttttttttttttttttt	2220
Db	30042	TTGTTAACAGAACTTCCAAATTTTTTTTTTTTTTTTATGCAACAAAGAAATPACATGAGAAA	2998
QY	2221	ctatttctgttgggaatggaagtagtaataatacatgaacaaacttlaaaaatatata	2280
Db	29982	CTATTGTGTTGGAATGGAAGTAGTAATATATACATTAAACCAATTTTAAAAATTTATATA	29922
QY	2281	agcctaatacgcgcctcaagaatgattatctagtagtgtataataatagcatggtgcgat	2340
Db	29922	AGCCTAATACGGGCTCAAGATGTTTCTATAGTAGGTATTAATTAATGACATGGTGCAT	2986
QY	2341	tcaagaattgggacacaaatgaaaaacgaattaaatatataacttlaaataataaataat	2400
Db	29862	TCAGATTTGGGACACAACTAATAAACGGAATTTAAATATTAACTTTTAAATTAATTAAT	2980
QY	2401	ltgagtaaatgtagtcttctgactacttggaggggcaaaaaagacaaatggccaagtctac	2460
Db	29802	TTGAGTAAATGTGTTTCTGACTATTTGAGGGGCAAAAAAGACATGCCAAAACTCTAC	2974
QY	2461	gggtttgactgtccagtttggtaataacttaaaactcgtcttggacgaagcgttgg	2520
Db	29742	GGGTTTGACTGTCCAGTTCGGTATTAATCAATAACTCTGCTTTGACCGCACGCTGCTG	2968
QY	2521	taggggtccctctbacaatttctacgttcttccacctacgttggacccaacctttccca	2580
Db	29682	TAGGGGTCTCTTGACATTTTCACTGTCTTACCCCTACGTCGACCCACACCTTTTCCCA	2962
QY	2581	tatccctaaggtaatttcttggaaatcccaattaaacgaatgtagacgtaaccgaactcc	2640
Db	29622	TATCTTAAGGGTAAATTTTGGAAATTCCTCAATTTTAAACGATTAAGACCGTACCGGCTTCC	2956
QY	2641	tggagatctcgttggagaatttacaanaattatagaacgaatgggtcttataatttaa	2700
Db	29562	TGGGATTTCTGCTGAGCATTTTATCAAAAATTTATTAGCACGAATGGGTTTATTATTAA	2950
QY	2701	aactccaacttgaatcgaataaaatttcaataaacaatttcaagatgatactcgtaagatc	2760
Db	29502	AACTCACAACCTGATCAGATTAATAATTTCAATAAACATTTTACGATGCAATTCGTAGCATCT	2944
QY	2761	atctcaatgacttttttttttcttccacacggttggatggaagtataatgactataatgacgag	2820
Db	29442	ATCTTAATGACTTTTTTTTTTTCTTCAACGCGGTGATGAATCTTAACTTATTATGACCAAG	2938
QY	2821	acaattcaattatagataatacctaataatcgaatataattatataataatgactttaa	2880

Dd	29382	ACAAATTTGATTTAGTATATATCCATTAAATCCATGATATTTATGATATTAATATAGCTTTAAA	29323
Qy	2881	ctatttcagcatcgcagctcttcctgcnaactttgttttcaattaaagattcaatataaa	2940
Dd	29322	CTATTTCCAGCATCCGAGCTTCTCGAACCTTTTGTTTAAATTAAGAGTTTAAATTAATA	29263
Qy	2941	aagattttaaagagcatalaagcagcacaanaagcaaatgaaacacggaagaaacaaagcca	3000
Dd	29262	AAGTATTTAAAGAGCACCATAACGAGCAACAAAGATTAATGAACCGGAGCAACAAAAGCCA	29203
Qy	3001	tgaagctcaatbgtta-gtttaagctttaaagaagattttaaactttaaacttaatacagat	3059
Dd	29202	TGAACCTCATTTGCTTAGGTTTAAGCTTAAATAGAATTTTATTAATTTAATACAGAT	29143
Qy	3060	gataacaattatatttctbgaacttctttaaagacccctcttaacaacagaaagctccctt	3119
Dd	29142	GATTAACAATTAATATTTTCTGACCTCTTTTAAAAACCCCTCTPACAAACAGAACCTCCCTTT	29083
Qy	3120	ttcagtagaagctcgcatctcccaactttaaagaacaagccatagaagaagaaagagagtg	3179
Dd	29082	TTCACTAGAAGTCCGATTCCTCCCATTTTAAAGCAACAGCCATTAGAAAGAAAGTGAATG	29023
Qy	3180	agagagagagagaaactagctccatlytttcogaacaagacatcaactctctacctcttc	3239
Dd	29022	AGAGAGAGAGAGAAACTAGCTCCATGTTTCGAACAGAGCATCATCTCTTACCTCTTC	28963
Qy	3240	ttctctcccatcggttttgctctctctctctcttgaatctctcttgaagagaagaata	3299
Dd	28962	TTCTCTCCCATCGCTTTTGCTCTCTCTCTCTCTCTCTGATCTCTTGAAAGAAACAAATA	28903
Qy	3300	gaaaaaacagatctcaatctaccctccgggtaaatccggtgycatcttcttgttgaacca	3359
Dd	28902	GAAAAACGAGATTCACTACTCTCCGGGTAATCCGGTTGGCATTTCTTGTTGTAACCA	28843
Qy	3360	tcggttatcttaaacccgttaacccgcacaacactcgtgtaactcatgcaacaactgtct	3419
Dd	28842	TGCGTATCTTAAACCGTAACCCGCAACACTCGTGAGCTTCAATGCAACAACTGCTCT	28783
Qy	3420	ccaagtaaacaacaacatcttccaaanaactcaaaaaataactctcgttttttgaatt	3479
Dd	28782	CCAATTAACAAACAATCTTCCAAAAACTCAAAAAAATAATCTCTGTTTGAATTT	28723
Qy	3480	tgaactaatgttgttattttaaaglatvgtaagatatataataagacgnaactgttttgaga	3539
Dd	28722	TGACATAAGTGTGTTATTTTAAACGCTAAGGTAAAGATATAAATATGAACTTGTGGAGA	28663
Qy	3540	accacacgatcgrataagctgcatgctgcgcgcgcgcgtatagatcatcattnaacaagaagag	3599
Dd	28662	ACCAACGATCGTATACGTGATGCTGATGCTGAGCTTATATGATTCATATTAACAAAGAAAGAG	28603
Qy	3600	gctcttgaatgtagttatcctcagaagataaagttggagatctcttggnaaatgtgcgatct	3659
Dd	28602	GCTCTTGAATAGTATTATCTTAAGTAAGTATAGTGGATCTTTGGGAAATGTCGATGCT	28543
Qy	3660	tgttcttggtgtgtagacatgcaagaagatatvggaagtatctgccttaacttcttaagtca	3719
Dd	28542	TGTTCTTGTTGGTGCATGCAATGAGATATGGAATGGAAGTATCTGCTTAACCTTTAAGTCA	28483
Qy	3720	cgcacgctctaaacacttctacttaagaagatgtagagacatcttgttggctctga	3779
Dd	28482	CGCAGCTCTTAACACTATTCTACTTAAGAATGTTGAGAGACATCTTGTGTGCTTGA	28423
Qy	3780	ttctcttgcaacaacactctatttctctctgctcaagaacgagccaaaagtttattt	3839
Dd	28422	TTCTTGCGCAACAAACTCTATTTTCTGCTCAACACGAGGCCAAAAGGTTTATTTT	28363
Qy	3840	tatctttaaatttgtaaaatttttttgtttaagaactttaagttcttaactttttt	3899
Dd	28362	TATCTTTTATTTTGCTAAATTTTTTTTGTTTATGAATCTTTGAGTTTCACTTTTTTTT	28303
Qy	3900	ttttaaactgaacagttacgtttaaactcaatgagcgaagacataatgagtaagtaacgtg	3959

Dd	28302	TTTTATTGAACACTTTTACGTTTAACTCTAATGCGGAAGCATATTAATGAGTATGCGATCTCG	28243
Qy	3960	gagaagaagaaacagagcaatlaaagaagaagtaagtaacttccatlyaaagagatgct	4019
Dd	28242	GAGAAAGAGAAACAGAGCAATTAAGAAGAGTATGTAACTTCATGAAAGAGTGTCT	28183
Qy	4020	ctgcccctcaatctccaggaactcttatataaagctctcgagtaactttt	4079
Dd	28182	CTGCGCTCTAAATCTCCAGGAACCTCTATCTATTAAGCTTCAGTACATTTATTTT	28123
Qy	4080	tttttcgctaaagtcacaactctcaatlaagttttttaaacttcttatgctgtaaa	4139
Dd	28122	TTTTTGTCTAAAGTCACAAACCTCATATATAGTTTTTATTTTATTTATGTGTAA	28063
Qy	4140	taaaatctcaaatgctgctgctgtaacagagcaagatgaaagctatgagagaa	4199
Dd	28062	TAAATATCTTAAATATGGTGTGTGTACAGCAACCATTTTGAAGTTCATTAAGAGAA	28003
Qy	4200	aatgaaagagaagaaattgatatacaaggaagaagatcaagaagaagaagtgaaac	4259
Dd	28002	AATGGAAGAGAAATTTGGATATCAAGGAAGATCAAGAAAGAAAGAACTGAAAC	27943
Qy	4260	agaggaagaaagcagagatgaaagtgatcatglttaagaaacaaagacagacgta	4319
Dd	27942	AGAGCATGAACAGAGATGATAGATGATCATGTATGAGAAACAAAGAACAGCATGA	27883
Qy	4320	ttcttgagatgaggttttgaaacttgaacttataagagaggaagcaatttcgaactc	4379
Dd	27882	TCTTTTGGATGGGTTTGTGAACATTCGAATTTATCGACGAGCAAAATTCATCTCAT	27823
Qy	4380	tcttaagcttgtaattgcgcgacaatgaaactcttcgtacatgctctgcatact	4439
Dd	27822	TCTTAGTTTGTATTTTGGCCGACATGAGACTTCTTCTGTAAGCCATGCTCTGCTATCTT	27763
Qy	4440	cttcttgaaagcttgcccttaaaagccgttgaaagccttaaggtaagataatataacagca	4499
Dd	27762	CTTCTTGGAACTTCGCCCTAAAGCCGTGAAGACCTTAAGGTAAAGATTAATAACGCA	27703
Qy	4500	caagtaattactacaaatgctacgataataaagattatagaactatctcatia	4559
Dd	27702	CAGGTTATTTACTACCAATTTGTTACGATATATATTAAGTTATTTATACAAATTTCTATT	27643
Qy	4560	gaatatcacatgtaaaaaagiatgatatlttaattgtccactaatlttatgtatc	4619
Dd	27642	GAAATTTACGAGCAAAAAAGTATGTATTTTAAATGTCCATCAATTTTATGTTTATTTT	27583
Qy	4620	ataactttgaaggaagacatcttgaaatcgcgagggccaaaggaagacagagagctca	4679
Dd	27582	ATACTTTTGAAGAGACCATCTTGAGTCCGAGGGCCAAAGAAAGAACTTAGAGATCA	27523
Qy	4680	gaatlaaatltggatgattacaagaanaatggaacttactcaatgctgatttactacat	4739
Dd	27522	GAAATTAATTTGGAGATTTACAAAAAAATGACATTTTACTCAATGTATGTACTATCAT	27463
Qy	4740	tctcatatttactatcatgctcatatgattatgaagaaccaaataatattgctttct	4799
Dd	27462	TCTCATTTTATTTATCTATGTTTCATATGATTTATATGATAAACCCAAATATTTGATTTTTT	27403
Qy	4800	tttgatgctgctgaaagctacaatgaaactctcgatctgggaagttaagtgtttt	4859
Dd	27402	TTTTGGTGTGTGAAGGTTATTAATTAAGAACTTTCGATTTGGGAATGTACTTATGAGTTTTT	27343
Qy	4860	gcatcgcaaaagcactcaaaagatgctcgltacaagaagtaaaactttacgtacaataatltt	4919
Dd	27342	GCATCGCAAGACACTCAAAAGATGTTCCGTACAAAGATTAACCTTTACGTCAAAATTTTTT	27283
Qy	4920	aaataagaactcggaaatctgaactcttatgttgaaanaaatataaanaataattaca	4979
Dd	27282	AAATATATTAATCCGATATTTGAAATCTTTTGTGATGAAGAAATATTTAAATATTTTACA	27223
Qy	4980	tttcttaatgcttg-aaaaaagatacagataccctcagtgagtgagaaagtgtaaccgctg	5038
Dd	27222	TTTTTATTTGTAAGAAAAAGATATAGATATCCCTAGTGGGAGAAAGTGTATCCGGTG	27163

QY 5039 atctcagcgctacatttgataatctcgtltatgaccaactaaatctttaaactctg 5098
|||||
Db 27162 ACTCAGCCGCTAATTGGATTAATCTCGTTATGACCAACCTAATCTTTAATCCTGG 27103
QY 5099 agatgcaacaggttaataaagttctctcgttactatgaaataatagttagtt 5158
|||||
Db 27102 AGATTGGCAACAGGTAAATAAAGTTCTCTCGTTACTATGCAAAAATTAAGTAAAGTT 27043
QY 5159 ttctcatctatcgtacgaatagatacgtcctaagtgatttaccatctatagata 5218
|||||
Db 27042 TTTTCAATCTATTGATGATAGATTAGTCTCTACGTGATTTACTATCTATAGTATATA 26983
QY 5219 cggagaacttaactcgtgcaaaaaacttattatattatcttccaagtlagactaa 5278
|||||
Db 26982 CGAGACCTATTATCTGCAAAAACTTTTATATTATCTTTCAGTTAGATCTTAA 26923
QY 5279 caagtcatagtatcattgtatccactgaagaacataataaataaagaagaagaag 5338
|||||
Db 26922 CACGTATGGATCTATTGATCACAATGAAGCATTAATAAATAAAGAGAGAAAGAG 26863
QY 5339 acgtgtgtgtatgaagtgatcgtgaagaacaattatagtaggtgtatgctttaa 5398
|||||
Db 26862 ACCTGTGTGTGTATGATGCTACGTGAACAATTAATAGAGATGGTATGCTTTAATGA 26803
QY 5399 cgttagagactgcctaaataatcttataatcgttgaccgtttattatagtcagact 5458
|||||
Db 26802 CGTAGAGCTGCGCTTAATATTCTTATATCGTACCGCTTATATTATAGCACGGCT 26743
QY 5459 ttgatcaaatltaagaatttgacaggaagatggtacacacgcttgaagatctcaacgc 5518
|||||
Db 26742 TTGATCAATATTAAAGATTGACGACGAGATGATACAGCGCTTGAAGGATCTCACGGC 26683
QY 5519 cgaatcgtgtacgtcgtttagatcttcgcacgcttgaactgtttataactatagat 5578
|||||
Db 26682 CGATGACTTGTACGTGCTGCTTAAATCTCTGCACGCTTACGTGTTTAAATGATTAAT 26623
QY 5579 aactctaatatatacaactatcaaatcgcggaatlagagaataataactatagat 5638
|||||
Db 26622 AACTCATTAATTAATTAACACTCAATCGCGAATAGAGAAATATCTAATATAGTAT 26563
QY 5639 tatatgattatattagaataacttacttgaataagaataaagaatagtagtactat 5698
|||||
Db 26562 TATTATGATTATATTATAGATAAATCTTATGAATAAATGAATGAATGATGATATG 26503
QY 5699 tatagtagtggggaaggaagtgatgagatgagatgaatgaacccaagctgtgtg 5758
|||||
Db 26502 TATAGTAGAGGGGGAAGATGAAGAGGTGAGAGATGATTAATGA -CCACGCGGTGG 26444
QY 5759 tgcacaagaacagtgcttcttcttcttcttcttcttcccaacttcttcttgggggtta 5818
|||||
Db 26443 TGCACAAGACAGTGTCTTCTCTCTCTCTCTCT -CCAACTCTTCTTCTTGGGGGTTTA 26385
QY 5819 ttctgattatataaatcgttctgttcttcttcttcttcttcttcttcttcttcttcttct 5878
|||||
Db 26384 TTGTGATCTATTAATAATCGGTTGTCTTTTCTTTTCTGACGACACCAAAACACGGAGC 26325
QY 5879 gtcactgctgaagaagtgtgtgttcttcttcttcttcttcttcttcttcttcttcttct 5938
|||||
Db 26324 GTGATCTCTCAGGAAGTGTGATTTTCTGACGTGGGAAACAACTACATGCGCTTGGAGG 26265
QY 5939 agggccaaggtctatgtctgtgtctcgaagctagccaagttagaagaatgcaagttcttct 5998
|||||
Db 26264 AGGGCCAAGGCTATGCTGTGCTGACAGGTAGCAAGTTAGAAATGGCAGTGTATTCTCA 26205
QY 5999 tcatctagttcttataatctaatgtggaattagagaagaatgacacacatttcttcttct 6058
|||||
Db 26204 TCATCTAGTTCTTAATTAATTAATGGCAATTAAGCAGAAATGAATTAACCAATTTGCTTTTCC 26145
QY 6059 ttctgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 6118
|||||
Db 26144 TTTTGTGATTTTCTTAACGGTTTGGCTTATAGGTTCTCGATCTCTGTAATAAAAAA 26085

QY 6119 aaagatgaagaatcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 6178
|||||
Db 26084 AAAGATGAAGATTTTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26025
QY 6179 gcccaatgataataaatttgataaataatattatgataatcgttcttcttcttcttct 6238
|||||
Db 26024 GCCCAATGATATTAATAAATTGGATTAATAATATTATGATATTCGTTTTTAACTTGG 25965
QY 6239 gttttagaagaaggttcttcttcttcttcttcttcttcttcttcttcttcttcttct 6298
|||||
Db 25964 GGTTTAGAAAGGTTTGACTTCTGAAAGTGCAGATGATATGATTTGGAGCTTAGG 25905
QY 6299 ttgagtccttggacatttgatgagatgltgtatgaltatagtgatgacactaaacc 6358
|||||
Db 25904 TTGAGCTTTGGACATTTTATTTGATGATGTTGATTAATAGTGCACATTAATTAACC 25845
QY 6359 ttaaatggcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 6418
|||||
Db 25844 TTAAATGGCTTTCTATTAAGGCCCAATTAATTAATGATTAATTAACAAAGTACAACTTTTA 25785
QY 6419 ctctgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 6478
|||||
Db 25784 CCGCTTTTGTGATCGAAGCAATTAACAAATTTGCAATATGCAAAACGCAAGATTAATGTA 25725
QY 6479 acactcgtgtgtcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 6538
|||||
Db 25724 ACACCTGTGTGTCTGATGAGGAAATCATTTGGCTGGAACATGCAACATGCAACAA 25665
QY 6539 aactcgtcaatattagatcacactccttcttcttcttcttcttcttcttcttcttct 6598
|||||
Db 25664 AACCTGTCAATTAATGATACCTCTATGACGTTTCCAACTTTATCTGATTTTAT 25605
QY 6599 cgtgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 6658
|||||
Db 25604 CGGTGTACATTTGACAAAGAGTGTAGTGTCAAAAGCATTAATGAATTAACATFACCTGT 25545
QY 6659 caggaataagaaggttcttcttcttcttcttcttcttcttcttcttcttcttcttct 6718
|||||
Db 25544 CAGGATTAAGAGGTTCTTAATTAATGATTTCTTTTATTTGAGAAAGAGAGAGGACCTTT 25485
QY 6719 tactgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 6778
|||||
Db 25484 TACTGTTTCTGGCTCTATCTCTTTAAATGAAGAGGTTTCTGTTTACTTCTTCTATCT 25425
QY 6779 catcatctttagatcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 6837
|||||
Db 25424 CATCATCTTAGATCTCTCTTCTGACGATTAAGATTAATCTGTTAACCAAGCAATGGGT 25365
QY 6838 ctcatcttctttagaaga -ggtcttcttcttcttcttcttcttcttcttcttcttctt 6888
|||||
Db 25364 CTCCTCTTTTGAAGACAGGCTTTTTCAAAGTCTGTAGTCAAGCCAAAGCTT 25313
RESULT 2
AF044216 4818 bp DNA PLN 06-MAR-1998
LOCUS Arabidopsis thaliana steroid 22-alpha-hydroxylase (DWF4) gene,
DEFINITION complete cds.
ACCESSION AF044216
VERSION AF044216.1 GI:2935341
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 4818)
Choe, S., Dilkes, B.P., Fujioka, S., Takatsuto, S., Sakurai, A. and
Feldmann, K.A.
The DWF4 gene of Arabidopsis encodes a cytochrome P450 that
mediates multiple 22 alpha hydroxylation steps in brassinosteroid
biosynthesis
JOURNAL Plant Cell 10 (1998) In press

REFERENCE 2 (bases 1 to 4818)
AUTHORS Choe,S., Dikes,B.P., Azpiroz,R. and Feldmann,K.A.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-1998) Plant Sciences, University of Arizona,
Tucson, AZ 85721, USA

FEATURES
source
1. 4818
/organism="Arabidopsis thaliana"
/strain="Ws-2"
/db_xref="taxon:3702"
/chromosome="III"
/map="8 CM from nga6"
990. .4398
/gene="DWF4"
/note="dwarf4"
/allele="dwf4-1; dwf4-2; dwf4-3; dwf4-4"
990. .1055
/gene="DWF4"
join(<1133. .1353,1434. .1758,1844. .1996,2095. .2409,
2562. .2654,2746. .2824,2931. .3040,3795. .4398)
/gene="DWF4"
/product="steroid 22-alpha-hydroxylase"
join(1133. .1353,1434. .1758,1844. .1996,2095. .2409,
2562. .2654,2746. .2824,2931. .3040,3795. .4040)
/gene="DWF4"
/function="brassinosteroid biosynthetic enzyme"
/note="member of the cytochrome P450 superfamily; CYP90B1"
/codon_start=1
/product="steroid 22-alpha-hydroxylase"
/protein_id="AAC05093.1"
/db_xref="GI:2935342"
/translation="MFEENHTLLPLLLPSLLSLFLFLKRRNRKTRFNLPPGKS
GMPFLETIGLKYPTATLGDPMOQHVSKYKITSNLEGEPTIVSADGLRPFILQ
NIEGRLEPCYRISGILGKSMWLYVGMHMRSTSLNLSHALLRTLLKDVHRT
TLFVDSMOONSIFTSKODEKKTFTNLAKHISMORGEETQLKREYTERKGVVS
APLNLPGTAYHKALQSRATLTKFERMERKLDIEEDDEEEVETEDAEKSKSDH
VRKQRTDDLLGNVLKHSNSTEQIDLILSLFAQHESSVALAIFELQKCPKAV
EELREHLEIARAKKEIGSELNMDYKKMFQCYINETLRGNVRFHRRALKDY
RYKGYDIPSGMKLPIVSAVHLNDRYDQNLNPMRMQONNGASSGSGSTGMN
NVMPPGGPRLCGSELAKLEMAVFTHHLVKNWELAEEDDFAPFPVDFPGLPR
VSKRL"

BASE COUNT 1566 a 727 c 888 g 1637 t
ORIGIN

Query Match 69.9%; Score 4818; DB 12; Length 4818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2071 gaattcaataatccatataatcttgaatcctaataatataatagtcacat 2130
|||||
Db 1 GAATTCAAATATCTCCATTAATATTTTGAATCTACAAATTTATATATAGTCAAT 60
QY 2131 acaaatgcatagaaagttccaaaaaaatttgttaacagaactccaaattttttt 2190
|||||
Db 61 AACAAATCATAGAAAGTTCCAAAAAATTTGTTAACAGAACTTCCAAATTTT 120
QY 2191 ttatgtgaaagaataaacaagatagaaactatttgttgtgaaatggaaatgtata 2250
|||||
Db 121 TTTATGGAACAAGAAATACAGATAGAAACTATTTTGTGTGAAATGGAATGTATATA 180
QY 2251 tacatgaagcaaatcttaaaaaatataataagctataagcgcccaagtatgtatcta 2310
|||||
Db 181 TACATTAAAGCAAAATTTAAAAAATTTATATAGCCTATACGGCTCAAAAGTATGTTACTTA 240
QY 2311 gtaagtgttaataatgcatggtgcattcagaattggagacaacaatgaaaaacggaat 2370
|||||
Db 241 GTAGGTGTAATTAATATGATGCTGCGAATTCAGAAATGGGACAACAATGAAAACGGAAT 300
QY 2371 taataatataacttaataataataaatttgatgaataatgtgtcttcctgactattgag 2430
|||||
Db 301 TAAAAATATTAACCTTTAAATAAATTTGATGAATGCTGTTCTGACATATGAGG 360

QY 2431 gcaaaaaaagaacatgcacaaagtctacaggttctgactgtccagttcgttaataact 2490
|||||
Db 361 GCGAAAAAAGAACAAATGCCAAAGTCTACGGGTTTGACTCTCCAGTTGGTAAATATCT 420
QY 2491 aataactgtcttcttgacacgacgtcgtgtgagggtcctctgacatttctactgttct 2550
|||||
Db 421 AATTAACCTGTCTTTGACACCGCACGCTGCTGAGGGGTCCTTGACATTTTCACTGTCTCT 480
QY 2551 accctactcgttgagccacccctttcccatatctctaaggtaatttggaaatcccaat 2610
|||||
Db 481 ACCCTACTGCTGAGCCACCCCTTTTCCATATCTTAAGGTAAATTTGGAAATCCCAAT 540
QY 2611 ttaaccgatgaaacggttacacgactccctgggattctcgttgagacattatcaaaaat 2670
|||||
Db 541 TTAACCGATTGACACCGTACCGGACATTCCTCGGATTCCTGTGAGCATTTATCAAAAAT 600
QY 2671 tattaagaagaatggttatttaattaaaaactcacaacttgaatcagataaaattcat 2730
|||||
Db 601 TATTAGACGAAATGGTTTATTATTAATAAANCTCACAACTTGATGATTAATTTTCAT 660
QY 2731 aaacactttacagatgattcgtacgatctatctaaatgacttttttttctaccaggt 2790
|||||
Db 661 AAACACTTTTACGATGATGATGCTGACGATCTATGCTTAATGACTTTTTTTTCTACACGGT 720
QY 2791 ggaatgaagtatatagtaactatacagcagaagaactgtatagatatatcatcaatcc 2850
|||||
Db 721 GGATGAAGAATGATATGCTACTATTACCAAGACAAATGATTTATATATATATCAATTAATCC 780
QY 2851 atgatatattatgataataatagctgttaactatttcaagcctgcagcttctgcacatt 2910
|||||
Db 781 ATGATATTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
QY 2911 ttgtttttaaatttaaggtttaaataaataaagatataaagaagagcctaaagaagcaca 2970
|||||
Db 841 TTGTTTAAATTAAGGTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
QY 2971 aagaatgaaacagaggaacaaagacatgaagctcaatggttgaattgaagcttaata 3030
|||||
Db 901 AAATGATTAAGAACACAGGAACAAAGAACGATGAAGCTCATTTGTTAAGTTAATA 960
QY 3031 agaaagtattatataatttaataatgaagatgaatacaattatattctgacttcttaaa 3090
|||||
Db 961 AGAAGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
QY 3091 accccctcttaacaaagaaactcccttttccagtagaagcagatcccaattaaag 3150
|||||
Db 1021 ACCCCCTCTTAACAACGAAGCTCCCTTTTCCAGTAGAAGTCCGATCCCAATCTTAAG 1080
QY 3151 acaaaagccattagaagaagaagtgagtgagagagagagaaactaagctccacgttcga 3210
|||||
Db 1081 ACAAGGCATTTAGAAAGGAAGGAAGTGAAGAGAGAGAGAGAAACTAGCTCCATGTTTGA 1140
QY 3211 aacagagcatcaatactctcttaactctctctctctccatcgacttctgctctctct 3270
|||||
Db 1141 AACAGAGCATATCTCTCTTACCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
QY 3271 ctctctgattctctcttgaaagaaagaataaagaacaaacagatctcaatcactccgggttaa 3330
|||||
Db 1201 CTTCTTGATTTCTCTTGAAGAAAGAAATGAAAAACAGATTCAATCTCTCCGGGTAA 1260
QY 3331 atccggttgccattctctgttgaaacacatcgtttatcttaaacacgttaacacgcgcaaac 3390
|||||
Db 1261 ATCCGGTGGGCATTTCTTGTGTAACCATGCTTATCTTAAACCGTACACGCCACACAC 1320
QY 3391 actcgtgacttcatgcaacacacatgctccaagttaaacacacacatctccaaaatc 3450
|||||
Db 1321 ACTCGGTGACTTCATGCAACAACATGCTCTCAAGTAAACAACATCTCCAAAATCTC 1380
QY 3451 aaaaaataaactcctcgttcttgaaatctgacaaatgctgttcaatttaaggataggt 3510
|||||
Db 1381 AAAAAAATTAATTCCTGCTTTTGAATTTGACTAAGTTGTTTATTTTACAGGTATAGGT 1440
QY 3511 aagatatatagatcgaactgttcttggaacacacgactcgtatcagctgactgactgact 3570

OY	5731	agaagatlaaagcccaagctggctgggccaagaagctgtctctccctttt	5790
Db	3661	AGATGATTAATGACCCACAGTGGTGTGGCCACAGCACAGCTGTTCTTCTCTTTT	3720
OY	5791	cttcccaactctcttttttgggggttatttgatttaataatcggttttgcgttttt	5850
Db	3721	CTTCCCAACTCTCTTTTGGGGTTTATGTGATTATAAATGCTGTTGCGTTGTTT	3780
OY	5851	tttggcagcgccgcaaaaaacggggcgctacgttcaggaaagtgttgattttccgagct	5910
Db	3781	TTTTGTGACGAGCAGCAAAACAAAGGAGCGCTCATCGCAGGAAGTGGTAGTTTTGAGAGT	3840
OY	5911	ggggaaacaactcaatccgctttggggggggccaagcgatagtctgttccaagctag	5970
Db	3841	GGGGAACACACTCATACCTCCCTTTGGAGGAGGGCCAAAGCTATGTCTGTTCACAGCTAG	3900
OY	5971	ccaagttagaatgycagtgtttattcatcatcatagttcttaaatlccaattgggaatag	6030
Db	3901	CCAAGTTAGAAATGGCAGTGTATTATTCATCATCTAGTTCCTTAATTCAAATGGGAAATTAG	3960
OY	6031	cagaagaatgctcaacacttgctcttcccttttgatttttccaaagtttgccttcta	6090
Db	3961	CAGAAAGTATCAACCTTGTCTTTCCTTTGTTGATTTCCTTAGCGTTGGCTATTA	4020
OY	6091	gggttctcgtatctgtlaaaaaaaaagaatgaaglatcttctctctctt	6150
Db	4021	GGGTTTCGCGATCTGTATAAAAAAAGATGAAGATTTTATTCCTCTTTT	4080
OY	6151	tttttgataattttaatcatcttttttgcccaatgatatataaaatttggataataa	6210
Db	4081	TTTTTGATTAATTTTAAATTCATTTTTTTTTCCCAATCATATATAAATAATTTGGATTAATAA	4140
OY	6211	tattattggatatcgttttctttagctcgggtttgaaaaaggttctgcacttcgaagt	6270
Db	4141	TATTATTGGAATTCGTTTTTGTAGTTCCGGTTTGAAGAAAGGTTTCCACTTTGCAAAAGT	4200
OY	6271	ggacgatgatataagatttggagctagtgtagtcttvgacatttgatttgaatgtyt	6330
Db	4201	GGAGAGATGATATATGATTGGGACCTGAGTGTGATGCTTTGGACATTGTGATTGGATGTGT	4260
OY	6331	tgattatttagtgctgaacacttaaaaccttaaaagggtcttctataaggccaattat	6390
Db	4261	TGATTATTAGTGCTGACACTATTAACCTTAATGGGCTTTCTATAGGCCCAATTATAT	4320
OY	6391	tacgatataaacaagatgacaaacttactcgttttltatccgaagaatataacaattg	6450
Db	4321	TACGATTATAACAAAGGACAACTTTTACTGCTTTTGTATCCGAAGCATATACAAATTG	4380
OY	6451	tcaaatccaaacaagaattatbtaaacactcgttgtagtctagcgggaaatcat	6510
Db	4381	TCAAAATACCAACCAAGAAATTATGTATAAACACTCGTGCTCTCATGGGAAATCATTTGG	4440
OY	6511	gctggagactgaataatcaagaacaagaacccgttaactatlatggataacccctcatag	6570
Db	4441	GCTGGAGACTGAACATATAGAACACCAAGAAACCTGTCAATTAATGGATTACACTCCTATGAC	4500
OY	6571	ggttccaacacttatcttgattcttctatcgtgttataattgacacaagaagtatagtgta	6630
Db	4501	GGTTCCAAACCTTATCTTATTTCTATTCGTGTTACATTTGACAAAGAGTTAGGTGTCA	4560
OY	6631	aaaggactaaatgatacaaatgctcctcaaggtaagaaagtctcataaattggtttctt	6690
Db	4561	AAAGACTAAATGTATACATATAGCTCTCGATGATTAAGAAAGTTCAATAAATGGTTCCTT	4620
OY	6691	attcttgaagaagaagaagagagacttcttactgttcttctgggttccatattccttlaaaga	6750
Db	4621	ATTCTTGAGAAAGAGAGAGAGAGCTTTTACTGTGTTCTTGGGCTATTCCTTTAAATGA	4680
OY	6751	gagggttctgtttctactcttctctactcatcatctttagagatcctctctctagaagata	6810
Db	4681	GAGGGTTCGTTTTTACTCTTCTATCTCATCTATCTTTTGAAGATCCCTTTTATAGAGAGATA	4740

QY	6811	aagtaatcctcttcacgaacgatgctcattcttbaagcaagctctttccaagctc	6870
Db	4741	AAGTAATCTCTGTTACCAAGCAATGCTCTCATCTTTGAAGACAGCTCTTTTCCAAGTCC	4800
QY	6871	tagtcaagccaagctt	6888
Db	4801	TAGTCAAGCCAAAGCTT	4818
RESULT	3		
LOCUS	AtF18022		
DEFINITION	Arabidopsis thaliana DNA chromosome 5, BAC clone F18022 (ESSA project).		
ACCESSION	AL163817		
KEYWORDS	AL163817.1	GI:7573446	
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota: Viridiplantae; Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots; Rosidae: eurosids II: Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE	1 (bases 1 to 100269)		
AUTHORS	Bevan, M., Murphy, G., Rudley, P., Hudson, S., Bancroft, I., Mewes, H.W., Rudd, S., Lemcke, K. and Mayer, K.F.X.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 100269)		
AUTHORS	EU Arabidopsis sequencing project.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-APR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de,Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@brc.ac.uk		
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ .		
FEATURES	Location/Qualifiers		
source	1..100269		
	/organism="Arabidopsis thaliana"		
	/variety="Columbia"		
	/db_xref="taxon:3702"		
	/chromosome="5"		
	2255..6118		
..gene	/gene="F18022_10"		
gene	complement(join(2255..2376,2496..2602,2699..2773,2896..3004,3094..3217,3419..3493,3582..3743,3871..4364..4393,4488..4553,4723..4759,4849..4922,5124..5180,5289..5339,5466..5531,5640..5775,5930..5981,6091..6118)/gene="F18022_10"		
exon	complement(2255..2376)		
	/gene="F18022_10"		
	/number=1		
CDS	complement(join(2255..2376,2496..2602,2699..2773,2896..3004,3094..3217,3419..3493,3582..3743,3871..4364..4393,4488..4553,4723..4759,4849..4922,5124..5180,5289..5339,5466..5531,5640..5775,5930..5981,6091..6118)/gene="F18022_10"		
	/note="strong similarity to protoporphyrinogen IX oxidase, Glycine max, AB025102"		
	/codon_start=1		
	/product="protoporphyrinogen oxidase-like protein"		
	/protein_id="CAB87761.1"		
	/db_xref="GI:7573447"		
	/translation="MASCAGVADHQIEAGVGRVAVGAGVSGLAAYATLKSGILNVTVFEAGRVGKGLRSYMWONGLIMDEGSAQNTLEAEPEVSGSLDDGLERKOOPISOKKRYIVRNGVYPLPTNPIELVTSSVLINSGNSTFTELEPELTKKRSKVSADSAAEVSSEFQRHRCQEVYDYLIDPFVCGTSAADPDSLSMKSHSPDLINSGSITVGAIRPKPAKGGKSRDTSFGTKSGRGSFSTRGGWQILIPDLCSLSHDEINLBSKVLTSTYNSGROENWLSQVSHNETQRQNHYPADPLCNVKEKKVMKGGQFPDLNPLPETINMPLSVLITTFTELEKSVPLRFGVGLIPSEKQKGFETGLTFLSSMMFPDRSPSDVHLVYTFIGGSRNOELAKSTDELKQGVTSIDLQRLGLVEEPEPVSNVHYVRKAFPLDYDSSYDSVMEADIK		

[illegible]

QY	2061	aatgcttgctgcatccaatcaccataac-----attcttggaaactcacaatt	2116
Db	74423	TATATTGCAATATATTTATTTTATTTTAAATTATTAATATATATATTTATTTAT	74482
QY	2117	aatacttcgtcaataacaatgcatacgaaagctccaaaaaaatttgytaacagaaact	2176
Db	74483	TATTTTAAATTAATTAATTAATTAATTAATTTAATTTAATTAATTAATTAATTAAT	74542
QY	2177	ccaactcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt	2236
Db	74543	ATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTA	74602
QY	2237	tggagctgcatatacaacttaagcacaatttctaaaaatataaagccatacgcgctca	2296
Db	74603	TATATTTATTAATTAATTAATTAATTAATTTATTTATTTTAAATTAATTAATTAATA	74662
QY	2297	aagtatgtatctcagtaggctgaacttaataatgcatctgcatcagaatcggagacaac	2356
Db	74663	ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTATTTT	74720
QY	2357	aatgaaacgcgaatataaacttaactttaaataataataaataattga	2404
Db	74721	ATTTTAAATTAATTAATTTATTTATTTCTTTAAATTAATTAATTAATTTTA	74768

RESULT	5
LOCUS	AC004157
DEFINITION	AC004157 169546 bp DNA HTG 12-AUG-2000 Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***, 2 unordered pieces.
ACCESSION	AC004157
VERSION	AC004157.8 GI:9797712
KEYWORDS	HTG; HTGS_PHASE1. malaria parasite P. falciparum.
SOURCE ORGANISM	Plasmodium falciparum Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE AUTHORS	1 (bases 1 to 169546) Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T., Kirdl,O.B., Conway,A.B. and Davis,R.W. Plasmodium falciparum 3D7 chromosome 12 Unpublished
JOURNAL TITLE	2 (bases 1 to 169546) Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W. Direct Submission
JOURNAL	Submitted (19-FEB-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT	On Aug 12, 2000 this sequence version replaced gi:8810447.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	23466:	contig of 23466 bp in length
*	23467	23466: gap of unknown length
*	23667	165946: contig of 145860 bp in length.

```

FEATURES
  source      Location/Qualifiers
1. 169546
   /organism="Plasmodium falciparum"
   /db_xref="taxon:5833"
   /chromosome="12"
   /clone="PRYAC293"
   /clone="3b7"
   /clone="1581 c 15705 g 68389 t 200 others"
BASE COUNT  69871 a 15381 c 15705 g 68389 t 200 others
ORIGIN

```

Query Match	1.9%	Score 133.2	DB 60	Length 169546
Best Local Similarity	4.7%	Pred. No. 2.1e-09		
Matches 1041; Conservative	0	Mismatches 1248	Indels 39	Gaps 12

D	85844	TATATACATTAATTAATATATATATTTTTTTTAAATAAAATAATTAATTAATTAAT	85785
Q	639	gagtagtgacccaatggagtagtactggaatactcaagaagataagagatgcaaccaa	698
D	85784	TATATAATTAATATATATATTAATATGTAATTAATTAATTAACAAATTAATTAATTAATA	85725
Q	699	aggcttggaaataatgctcccgccagcctctctccacaatcaatgcagccctatgg	758
D	85724	AATATATTAATTAATTAATTAATTAATTAATTTATTTATATTAATATATTAATTAAT	85655
Q	759	attctcggaatctcgtaaaatctgcgaataacgctggaataatattactgttag	818
D	85664	TAAATTTAAATTTAAATTTAAGA-----AAGAAATAATTTAACTTTTA	85623
Q	819	ctgactccaatactagctcccgagatctgcgaatactctctcttaagcatattgctc	878
D	85622	TTTAAATTAATTAATTAATTAATTAATTTTATGTTATT-TATTAATTAATTAATTAATA	85564
Q	879	tctctttctgtcgtctctctcaactatactcgcggaatataatgaataacatgata	938
D	85563	TTTTATTTTATTTATTTATTAATAATAATTAATTTTAAATAATATTTATTTAAATTTAA	85504
Q	939	tatccaaacaatctgctcgggacatttgataaactcttctccaacatctcggac	998
D	85503	ATATAAACATTAATTAATTAATTAATAATTAATTAATTTTAAATTAATAATTAAT	85444
Q	999	actggactcgaccttaaaatacgaattctacgcgcactagctagattactacataa	1058
D	85443	ATATTTACATTTTATTAATTAATAATAATAAACCATTTATTAATTAACCTTAATTAATAT	85384
Q	1059	agcataaaggaccgcttcaagctatttacaagaatcaaacctgaataatagcttgaat	1118
D	85383	AAAAATAAATTAATTAATTTGTGTAATTAATTAATTAATAATAACATTTATTAATAATTAAT	85324
Q	1119	cccttagaatacttggaaatccgctgtctgaataataagatttagtggtaacaaa	1178
D	85323	TATATATATTAATTAATTAATTAATTAATTAATTAATTTTATTAATTAATTAATTAAT	85264
Q	1179	ta-----tgttaacatctagctggtccaacataacataatctctcagaaaaacaaact	1234
D	85263	AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	85204
Q	1235	taagaagaatcacatccatactggttagcctaaccttccgctgatactgatacag	1294
D	85203	TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	85144
Q	1295	agactaaaga-----tagttaatgctgctgaataatgaaatcacacgcgttgtaat	1348
D	85143	ATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	85084
Q	1349	aacttcggagacgctgctgtagcaatcaacgcaaatcaatctctgtgtgtaacaataa	1408
D	85083	TTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	85024
Q	1409	aacaaacaagaagaaagaacgaattctctggaatccatccaatgactcaaatg	1468
D	85023	ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	84964
Q	1469	catgatactttcgggtacagttctcgaagctcctcacaagcgtgaacacatctgcaacta	1528
D	84963	TTTATTAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	84904
Q	1529	ttaaatctgctcttcaatgcaatcttcaacatattatctagcttagctgaatttaaga	1588
D	84903	ATATATATTTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	84844
Q	1589	gcgaactctgtaacatcacataattatatactagatactgatactgatactattccaataca	1648
D	84843	TATATATTAATTTAATTTAATTAATTAATTAATTAATTAATTAATTTTATTAATTAAC	84784
Q	1649	tactcttgatgcttaactaaatctctgtctctcctaaggtataataatcatctgag	1708

[illegible]

```
/organism="Plasmodium falciparum"
/db_xref="taxon:5833"
/chromosome="2"
/complement(1570..2424)
/gene="PF0490c"
/complement(1570..2424)
/gene="PF0490c"
/feature="predicted by GlimmerM"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC71887.1"
/db_xref="GI:3845198"
/translation="MKENKIMDYLCPLDVPVREKSGKSLKSSSTKSDYK
SIFSKRDSHKKSGSPRSGEINRSGSPKPYNNRILINKNNYNGRPHNR
DNKGRGSGSPDRSGSPKRFISNNKSSVSKRGNGSKMGKSPKAPTSR
TYVTKRLNNKYTSAPYKKNINISLXKRNRPALNTKSKPGVTKSSVPKRIK
GLKGSLSKSTKSTSGSKRPLKIFLSKIKIVTSLNKIPSLKQKNEVNLPEL
NNAVTNKN"
gene
11241..14606
/gene="PF0495w"
11241..14606
/gene="PF0495w"
/feature="predicted by GlimmerM"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC71888.1"
/db_xref="GI:3845199"
/translation="MPALKKNTYREGVNICFSYLLKLYLKSNEYVNLNENR
TNKIKKSKAQSLFDGLNIDHLLIFKLPYKCKAYECISKEVYKLLDEKRC
FNYISLDDIIOQVKEIDELDEFTDYNEYLEVKNIDNVLNKINEIFYNKDTEHNR
ELIGKICNKMISYIHEMNGNELIHLIYFPRMKNIDNLIIFYNYNYVDHLYFN
HEIYKLIFFNKYLNNSNIPEKNLLOEMFNLYYREIKENYLIKNNKEIYK
CEAFHEVNDHIDNEKILNLIPLVDNSLIDIDNNKLNNNLNINEIYISKL
NFCYTLKKKKTIDNDMTYKLEVIKATHTHIDCKTKNETFCSIDYSTLSLNSLNK
FLNKIIDKNEILYECLEKILNLIKRVNQSCLISLKNLYNLNNVYLVNV
LFDIMKFSLYLCNIFAGRIKTEENAVLIHNDQTNYSNENIDIIYOKRIEY
LPIKMEYKDFHEKIDSDLSIKLSNTEVKINEVNSDEYLFNNISCLINFLY
NRSVKKYKDYIYLINDISFYVKIKNNRTRKKKKFELSSMKELICKNNKSVL
RYIKLHEEDNPOKOYVCSLFIANLIFPKDILHPIYINMCHYKYKYNYECNLT
INEDIISLLTCSKROYFLENNSDRYCKRELHILKNIIDILKATLWYISID
NISIKFISLNSKYTCVENNLLLESLOSFEVATYKSGGIMHMDNNLNNNSE
KYBHYRILEYKKNLFIANKLIECLINFLYKKKYLYLVKQICPINLENL
KLILYANNLNYEMYGCEMLERVLSHKEONLFYNNKVEHMFKICHISE
DDYIEMNTMYLVLYDKLINSEBOSNLRNNSTNRFDEIKERKKYNNNTLIH
NNYKINERKSNNSNLSNLDKDKNNHNVEMDLDNKNENKIOEKONGNCN
CKDVLVDIINIFGFLKMEKKELFQIYVLCNITKFKRIVSSSLFHMDFKIIK
DMNKYICLENYKIKNECAFVYIDIVLEKER"
BASE COUNT      6284 a      1019 c      1106 g      6458 t
ORIGIN
```

```
Query Match      1.9% Score 129.4; DB 4; Length 14867;
Best Local Similarity 43.8% Pred. No. 9e-09;
Matches 1036; Conservative 0; Mismatches 1306; Indels 25; Gaps 10;
```

```
Dh 7392 ATAAATATATTAATAAACAATAATATATTAATTAATTAATTAATTAATTA 7451
Oy 406 ttgtgtcgtgtagttagttagttagttagttagttagttagttagttagttagt 465
Dh 7452 TATATTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7511
Oy 466 tgaagatttctctctctctctctctctctctctctctctctctctctctctct 520
Dh 7512 CACGTGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7571
Oy 521 aaatacagagtgatgcgcgttagttagttagttagttagttagttagttagt 580
Dh 7572 TAAATTAATAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 7631
Oy 581 aatccagaatcttcaagaagtagttagttagttagttagttagttagttagttag 640
Dh 7632 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7691
Oy 641 gtagtgaccatgagtagttagttagttagttagttagttagttagttagttagt 700
Dh 7692 AATATATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7750
Oy 701 gtcgtggaataatgcgcgttagttagttagttagttagttagttagttagttag 760
Dh 7751 TATTTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7810
Oy 761 ttctcgatctgtagttagttagttagttagttagttagttagttagttagttag 820
Dh 7811 ATAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7870
Oy 821 gatcacaatctctctctctctctctctctctctctctctctctctctctctct 877
Dh 7871 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7930
Oy 878 ctctctctctctctctctctctctctctctctctctctctctctctctctctct 937
Dh 7931 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7990
Oy 938 atat-cacaaacaatctctctctctctctctctctctctctctctctctctct 996
Dh 7991 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8050
Oy 997 aacatgcgacgcaccccttaaaatagatctctctctctctctctctctctct 1056
Dh 8051 ATAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8110
Oy 1057 aaagcataaagacccgctcaagctatctatctatctatctatctatctatct 1116
Dh 8111 AATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8167
Oy 1117 atcccttagaataatcttggaattaccggtctgttagttagttagttagttag 1176
Dh 8168 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8227
Oy 1177 aatatgtaacatctagtggtcaacatcatcatctctctctctctctctctct 1236
Dh 8228 ATAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8287
Oy 1237 agagaagtaacatcatcatatagtggttagttagttagttagttagttagt 1296
Dh 8288 ATTAATAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8347
Oy 1297 actaagaagtagttagtgtagttagttagttagttagttagttagttagttag 1356
Dh 8348 TCTATTTATTT-TTATTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 8405
Oy 1357 gacgtagttagttagttagttagttagttagttagttagttagttagttagttag 1416
Dh 8406 TTGTTTCTTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8465
Oy 1417 caagaataaagaagaagatctctctctctctctctctctctctctctctctct 1476
Dh 8466 TTAATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8525
```


[illegible]

DEFINITION	Genomic sequence for Arabidopsis thaliana BAC T12C24 from chromosome I, complete sequence.
ACCESSION	AC025417
VERSION	AC025417.4
KEYWORDS	GI:9438236
SOURCE	HTG.
ORGANISM	thale cress. Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
AUTHORS	1 (bases 1 to 106142) Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shin,P., Altafi,H., Bel,Q., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howng,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Federspiel,N.A., Theologis,A. and Ecker,J.R.
JOURNAL	Genomic sequence for Arabidopsis thaliana BAC T12C24 from chromosome I
REFERENCE	2 (bases 1 to 106142) Ecker,J.R.
AUTHORS	Direct Submission
JOURNAL	Submitted (09-MAR-2000) Arabidopsids thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE	3 (bases 1 to 106142) Ecker,J.R.
AUTHORS	Direct Submission
JOURNAL	Submitted (23-JUL-2000) Arabidopsids thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE	4 (bases 1 to 106142) Cheuk,R., Shin,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howng,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
AUTHORS	Direct Submission
JOURNAL	Submitted (27-JUL-2000) Arabidopsids thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
REFERENCE	5 (bases 1 to 106142) Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howng,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Federspiel,N., Theologis,A. and Ecker,J.
AUTHORS	Direct Submission
JOURNAL	Submitted (09-JAN-2001) Arabidopsids thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
COMMENT	On Jul 25, 2000 this sequence version replaced gi:17940297.
FEATURES	Location/Qualifiers
SOURCE	1..106142 /organism="Arabidopsids thaliana" /db_xref="taxon:3702" /chromosome="1" /clone="T12C24" 121..786 /note="similar to CLV1 receptor kinase like protein emb1/CA18252.1" /codon_start=1 /evidence="not experimental" /product="T12C24.1" /protein_id="AAP88073.1" /db_xref="GI:9502366" /translation="MQLTSEFVNGSLYDNLHLRPGTSSSYSGNTDLMWHRRFQIA
CDS	

CDS
 LSTAKALSFLLHNDCKPAIHLNVKSTNILLDEREAKLSQYGLEKEFLPVMDSEGLTRK
 FHNAGVYIAPELAQOSILASSEKCDVYSYGVLLLELYTGRKPEPSPENQVLIRDYVR
 DLLETGSASDCPDRRLREFRENELLIOVMKIGLICTSENPFLKRSMAEVOVLIESIRNG
 FGS"
 join(1701..1863,1982..2194,2289..2428,2519..2614,
 2962..3103,3181..3226,3339..3483,3599..3742,3849..3968,
 4112..4168,4249..4308,4456..4563,4722..4864,5060..5138,
 5247..5309,5402..5532,5618..5723,5908..6096,6220..6495,
 6595..6750,6873..7034,7168..7284,7367..7543)
 /note="unknown protein; similar to ESTs gb|A199904.1,
 and dbj|AV547742.1"

/codon_start=1
 /evidence=not_experimental
 /product="T12C24.2"
 /protein_id="AAf88074.1"
 /db_xref="GI:9502367"

CDS
 /translation="MDGROVESVDLLEKATNRKNTCTMAAGNDVYLGTSKWTI
 RYDFGVSSNDIDLAIVGTEQSIHKEVDPGSHCATVGVGAETFTAKMLKP
 RYLSRKLGLVAVANNRQITTEVSTRKILLQDQLEMAVDEKREKYLKLE
 LEELPAFKALQMETANISSGMRYVAVAPETLSPGTIGTLESVASYKERAHEM
 ELPEIPIRNOVLIDRLIRIKETKONSLLTWGITGHLNGAHSYKNDENV
 ENKALLDYSKLSGTRAVRGSMALESEYHLLIKKVKYVNISEDIIEEDPDITS
 DVSRSKIGLSDASANVFAYDQNSIFQVVIDEGDMKVLIDLKVAALANDC
 PLORDVYLVQASAFDKEYLKASRYANVYISEEYTLKESINEBALRTEL
 HKLDNLSKDKCQITMISTWATELYLDKINRLLEDLAIENRDESHSYIOEFRA
 SDCKDELDEATVILLESYGVRELVEYFANLKOYEIVLHYIOGGAKKALFLOKS
 SVSELOYOPAPELIMDAEYVESMANKNLPRLLITAMRYSGGRAKNTHEVI
 KYLECYVRLHNDPCGIIHLLSLYAKOEDGALLRLOCKFEKRGNGEPFETDRI
 ALRLCLKERTKACVHIYSMSMHEEVAALQCAESVDHLKTVSLQIDPELAAME
 ADVVEDEDELRKRLMLVAKHVKVQEGAKRENIKRAIALKEDGLIEDILPEP
 DFALIDPFKALSSLEDYKQIEOLKEEMNDATRGANDIRNDISALTQRYAVIDRE
 ECGVCKRIILMSGDFRMAQYSGAGLAPFVPGSHFACOLITHVSCAHEBA
 EHLLDLOKOLLGSETRRDINGNSDEPITSTTADKLRLSELDALIASCPGCELM
 INITLPIKPEBSQYSTSDLSERTLANORTISLPV"
 complement(join(7769..8120,8198..9009,9085..9591))
 /note="hypothetical protein"
 /codon_start=1
 /evidence=not_experimental
 /product="T12C24.3"
 /protein_id="AAf88102.1"
 /db_xref="GI:9502395"

CDS
 /translation="MERKQSNASTFADINEVEDEAEOLQOENNNKRFSGNREGN
 RGKORPFRGSRQVSLSTGFSVLNRESREDDKSLPSRSGRFGGEGGIIINGDRG
 KTDPSMERTKSTLSKOKSLPSIIREDIDENSLRTEDGFKDSINENYAGRFPAL
 RGPELDEVKNEDILLPKEOMPELPLRFGICGICGLSSOAVMLKASPAENTL
 HIRPLINLVYVESLAVYVSFPTYLKICFPEAVYREYHPVRVREPRAPVYCNF
 LAISVPMESPNNRYLHPATVCFPMGPFLELKITGOMISGCKRLCKAANFSSHLS
 VVNGFVGAIIASVVGWDEVAKFLMAGFAHVLVFLYQLRPLTSEALPELHPVYSM
 FIAAPSASIAAMNTIYGOFGDSRTECFIALPYISLVARINFTGFKFESVAMWSYF
 PMTASVATIKVAEAVGYPSSALALILFISITAMCVLFEVSTLLHAFVQTLPEPDL
 AIAITRKRLIREKKPFKRAYDLKRWTKQALAKISAKDFEAEEESH"
 join(11098..11279,11308..12211)
 /note="hypothetical protein"
 /codon_start=1
 /evidence=not_experimental
 /product="T12C24.4"
 /protein_id="AAf88075.1"
 /db_xref="GI:9502368"

CDS
 /translation="MSKKQRCCLMESIPHEVERILEKTCGGSALIHGICETMEIID
 RISILPKRTIDIASAIRSDYFNPDILKHTTLVLGSSSSVITPTPEKKEKEKE
 KEDEPELVSFSDGDLVCLYKWKSGVYVPTLRWRLPLSLQOLLISLGISVE
 LGYVCDIGGKDKITGTYKPVMLYNSLEIGLENATCEVDNTNMRVINSAPYR
 EETKILSFEDHTEFRVYKAPPTNVAKPIENGLNRCICVSKNNRNVYISFNGS
 NKTMRKESINIDVTSMRGNHTAAWPLALFYEKKKKKLLFPCRRASRTLWYDDE
 TESDYAFNDYSTGPLYCTOSLISIS"
 complement(join(13884..14297,14628..15299))
 /note="similar to phosphate/phosphoenolpyruvate
 translocator precursor gb|AAB0647.1; similar to ESTs
 dbj|AV524415.1, gb|A1995339.1, and dbj|AV542831.1"
 /codon_start=1
 /evidence=not_experimental
 /product="T12C24.5"
 /protein_id="AAf88101.1"

CDS
 /db_xref="GI:9502394"
 /translation="MVEAGSWTTRMSNPREDAAATPTIVDIPGPHSSASSPLK
 PFLSSPTVPTLITLIIAMGNSNGVLLKRLYLFYGGPRPIPLMTNMLSCAA
 YSSAVINAGIVPROHLTSROBLKILSIAFICLSVCGNTSLRATPVSENOAGT
 TPFTTAVSFLLTCKTESTEVYLLALPVSIGVILASSESEFHLFGLICVASTAGA
 LKSVOGIIILTSESEKLSHNNLLYMAPACILLPETLYTEGNVLKLEKATPDL
 IIFLNGALNVAALVNTNLFVLRKHTSALTLOLVGNGKAAVAGVSLIFRNPVTWG
 IAFEGVIMVAVYSEARSKSLNOK"
 complement(join(17619..17885,17959..18182,18485..18619,
 18717..18789,18871..18897,19006..19242))
 /note="similar to Cu/Zn-superoxide dismutase copper
 chaperone precursor gb|A052685.1; similar to ESTs
 dbj|AV442631.1, gb|BE038022.1, dbj|AV441271.1,
 dbj|AV530707.1, dbj|AV561064.1, and dbj|AV561353.1"
 /codon_start=1
 /evidence=not_experimental
 /product="T12C24.6"
 /protein_id="AAf88100.1"
 /db_xref="GI:9502393"

CDS
 /translation="MASILRSVATTSVAASAIPAIAPFSSSSSSSTNPQSOLN
 FSLRSRSPILLGSRFSVPAPATATSPRNHQEDRAPOLLTEPMVDMTEGYN
 AVNKRLEITIGIEKEVEDLSNOVRIIGSSPVKAMTQALFOIGKARLIGOGPDL
 VSAVAFAFKPDIFGVVRFQAVSMELARIEANFTGLSPGTHSCINIEYGLDTGAAS
 GSYLNPQDGTPEPLDGLTLEADKNGEAFYSGKKELKAVADLIGRAVVYRTDNRK
 SGGLRTAAVILARAGVGENYKILCSCDGYIWEATNDFVASKV"
 complement(join(19464..19818,19915..20047,20203..20316,
 20580..20642,20957..21104))
 /note="hypothetical protein"
 /codon_start=1
 /evidence=not_experimental
 /product="T12C24.7"
 /protein_id="AAf88099.1"
 /db_xref="GI:9502392"

CDS
 /translation="MARKISKEVEVLEKDDGFDRRLKILIRLKNVPSLSPSS
 SHCISLQYIWLVCDFKELGREVSLVNMWDLILLIGIVFIVEESLYLSIEMA
 LSLAQALVTLOEDLRNNMISVYKESLIRHGQNNKRTLOLSDAIEEYKMSLSOL
 SDGLMGIIIRSEDKMNEIRETVQSVYSLNLNPREKGTGAREMHEKIPPKREARD
 FMTSPASKOEOELKGAVEDNKGACGDVODNNGEAYNDEEDPELPPFG"
 join(23302..23622,23752..24150)
 /note="hypothetical protein"
 /codon_start=1
 /evidence=not_experimental
 /product="T12C24.8"
 /protein_id="AAf88076.1"
 /db_xref="GI:9502369"

CDS
 /translation="MNPDSILFPNAYENDLDFSLTPSPRVROEKPQNPVTH
 SAGIENDRONCETMTLSEIMKDDPEKKNRAHKLEROERONTSLKILTYLP
 SQYIKGRSSADHLEAVNYIKDQRIKEVSEKRDRIKSIITHPSRGEFSIRLSAS
 STCSGVDTNIAVVRPCLIGLEIVVSCNKHESCLSSVQLLAEOECFNVISCISTR
 LHOGFIITIASEVIYVLSI"
 join(24991..25437,25649..26173)
 /note="similar to D-3-phosphoglycerate dehydrogenase
 sp|P35136|SERB_BACSV; similar to EST gb|T245682.1"

Query Match 1.9%; Score 128; DB 12; Length 106142;
 Best Local Similarity 50.0%; Pred. No. 1,18-08;
 Matches 402; Conservative 0; Mismatches 395; Indels 7; Gaps 3;

Qy	3291	gaagaataagaanaaacagatcaatcaatccggtgaataatccggttgccattcttg	3350
Db	92868	GGAGAAATCCCTAAATGACAGGAAATCCACCTGTCATGGTGTCCCTTACCG	92927
Qy	3351	gtgaacacatcggtatcttaaacgcgcacgcgaacacatcggttgatcttaacgaac	3410
Db	92928	GCGAAGTATCAATTCCTCAAGCCAAACAAACTTCAGACATCCCTCTTTATCAAG	92967
Qy	3411	aacatgltcccaagtaaacacaacatcttcaaa--aactcaaaaataataatcctcg	3468
Db	92988	AGAGAGTTAAGAGAGTATCTCATCTCTGCATTCAAACAAATGTGATTAAATGCTTAT	93047
Qy	3469	ttttgaatttgactaatgtgtttatttaacagtgatgtaagatatataatgaac	3528
Db	93048	AATGTGTTTAAAGTGATGTGATATGTGCAGGTATGCCCAATTTTCAAGACCAAT	93107

YDLKMHKEINRLKLYKKRAFHLYITVADPQDILINNYKHIHKKHKKHIIHILHKKCKKE
IDLVKPKDLYCYCHMINNINNIFFSYIHVADNECYKTRFKAYNKLILFYNTLKNCKEYI
YKKRIKLYTLYLNNNISNDLPYINISYKNIYONKKNRKNRSHIFYSKANDLHLLF
LVTORLONCDLFSYLYKKRYENKNEKPNLYLHLYIVYSRKKKKKFFSITSSP
YSDVMDNCKDLSYFFPFSLYLFLIFLHYHDMVASYIFKTLFYSGLPYHYSLYNNI
MVYCGPKT"
join(7669, .7717,7800, .7829,7912, .7940,8064, .8106,
8302, .8368,8568, .8641,8713, .8812,8924, .9006,9122, .9258,
9369, .9505,9613, .9838)
/gene="PFC0581w"
join(7669, .7717,7800, .7829,7912, .7940,8064, .8106,
8302, .8368,8568, .8641,8713, .8812,8924, .9006,9122, .9258,
9369, .9505,9613, .9838)
/gene="PFC0581w"
/note="PFC0581w, hypothetical protein, len: 324 aa,
revised: new gene prediction, splicing prediction very
tentative"
/codon_start=1
/product="hypothetical protein, PFC0581w"
/protein_id="CAB90285.1"
/db_xref="GI:7711066"
/translation="MKKKHYSIKFVAVDKKYYVLCSDYRIIRINMGHIFRL
OKIPIHYFANEOLASSFESHKRYRTKEDIIDIGCKPNTIDYLSIESIKDDNPS
NOKITVTYAMKKNMDQLTPTYSNKRIDVYNASEFQWASEYIFLNIKSHRSPRA
LTKKDEIYISKNNFSPALSNDNSVTKRYIDVLLDNLNIISEETKPNFASGVGVY
TLYKKEKLTIMNRLILSKERYPNQVMDMKERIHITLFTYINLFFLSISHRYHDSV
ONLREKKNNSDKLDQDIDEDDEKYPDEILRAKKKSEEDDDEL"
7790, .7799
/note="potential splice acceptor sequence for exon 2 of
PFC0581w"
7830, .7835
/note="potential splice donor sequence for exon 2 of
PFC0581w"
7901, .7911
/note="potential splice acceptor sequence for exon 3 of
PFC0581w"
7941, .7947
/note="potential splice donor sequence for exon 3 of
PFC0581w"
8055, .8063
/note="potential splice acceptor sequence for exon 4 of
PFC0581w"
8107, .8112
/note="potential splice donor sequence for exon 4 of
PFC0581w"
8293, .8301
/note="potential splice acceptor sequence for exon 5 of
PFC0581w"
8369, .8374
/note="potential splice donor sequence for exon 5 of
PFC0581w"
8559, .8567
/note="potential splice acceptor sequence for exon 6 of
PFC0581w"
8642, .8647
/note="potential splice donor sequence for exon 6 of
PFC0581w"
8704, .8712
/note="potential splice acceptor sequence for exon 7 of
PFC0581w"
8813, .8818
/note="potential splice donor sequence for exon 7 of
PFC0581w"
8913, .8923
/note="potential splice acceptor sequence for exon 8 of
PFC0581w"
9007, .9012
/note="potential splice donor sequence for exon 8 of
PFC0581w"
9112, .9121
/note="potential splice acceptor sequence for exon 9 of
PFC0581w"
9259, .9264

Query Match	Best Local Similarity	Score	127 8:	DB 96:	Length 86827:
Matches 1051:	Conservative	0;	Pred.Nc. 1.2e-08;	Mismatches 1307;	Indels 29; Gaps 11
misc_feature	/note="potential splice donor sequence for exon 9 of PFC0581w"	9361..9368			
misc_feature	/note="potential splice acceptor sequence for exon 10 of PFC0581w"	9506..9511			
misc_feature	/note="potential splice donor sequence for exon 10 of PFC0581w"	9601..9612			
gene	/note="potential splice acceptor sequence for exon 11 of PFC0581w"	complement(join(10024..10044,10320..10336,10397..10445,10613..10674,10773..10798,10929..10999,11157..11237,11453..11515,11715..11767,11950..12040))			
CDS	/gene="PFC0582c"	complement(join(10024..10044,10320..10336,10397..10445,10613..10674,10773..10798,10929..10999,11157..11237,11453..11515,11715..11767,11950..12040))			
	/gene="PFC0582c"	/note="PFC0582c, hypothetical protein, len: 177 aa, revised: new gene prediction"			
	/codon_start=1				
	/product="hypothetical protein, PFC0582c"				
	/protein_id="CAB90286.1"				
	/db_xref="GI:7711067"				
QY	35	tacaaataaaatttcggtgttcggtatccggtacacagatcttggtgttcacaa	94		
Db	38839	TAAATATACAAATATTAATCATGCTTTTAAATAAATAATATATATTTTATTAATATATAA	38898		
QY	95	tttgatcccaagaataaataatataaattcatttaaaatatttacaagtaaat	154		
Db	38899	TATTAAGTATATATTTATATGTTATATATACATTTTATTTTAACTTCTTAATTTAAATATGG	38958		
QY	155	atccttacatgcatctgttatatacaaaaacatcatcctctgttatagagaataatcgag	214		
Db	38959	ATTAACATATTTTATTAATATTAATTTTACACATTTTATTAATAAAAAACATATATTTATATAT	39018		
QY	215	tttggaatttataaataaagaataacgattccatttgggttgattacacagtt	274		
Db	39019	AATTTATATATTTTAAATAAATAAATATTAATTTTAAATTAATTAATAAATAAATATATGTA	39078		
QY	275	gttttgggttccttcttggttatatgataatgagtaacaaagaagatctgtaagt	334		
Db	39079	ATAGTTATATATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	39138		
QY	335	gtaaacataatcgttatgacccccaataaataaataaataaacaacaacacccccc	394		
Db	39139	ATATATATATTAATTAATAAATAAATAAACAATAAATAATTAATTTTAAATTAATTAATAA	39198		
QY	395	cccgatatagtttgggttcggttgaagttatttgatcacataatcaatgcacatcc	454		
Db	39199	TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	39258		
QY	455	tttgattacatgaagattctctcccaatbaaaatttcgaattcatctcttgatt	514		
Db	39259	AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	39315		
QY	515	taaatataacgagtggaatat-ccgtttatcgatcaccccaatcatgatatagattc	573		
Db	39316	TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	39375		
QY	574	tgtgtcctacacgaacattatcaacagagatcttgagaataacacgaataaagaagaag	633		
Db	39376	ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	39435		
QY	634	ggaagaagtagtgcaccatcgagatctgtgataatcatcaagaagataagaagtacaa	693		
Db	39436	TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	39495		

[illegible]

Oy	1772	tgaaatgtaacccttact-----catgatgaattacccagatcttttcgcata	1827
Dd	40567	ATTATATTTAATTAAATTATTTATATATATAATTTATTTATTTATTTTAA	4062
Oy	1828	tgcagcctcaccaactgta-----tagtcaalcttttcgcaaatlaaattagg	1880
Dd	40627	TTAATTAATAATATATTTAATTTAATTAATTAATTAATTAATTAATTTA	4068
Oy	1881	aattcaatgcacatcataagaagaacaagctcgagttaactttaatttaaagcaa	1940
Dd	40687	ATTAAATTAATATATTAAATTTAATTATATTAATTTAATTTAATTTAATA	4074
Oy	1941	aatttcggaaaatgtataatactccaacaatatlaataaatacgcacctaaigta	2000
Dd	40747	ATATATTAATTAATTAAATTATATATATTTATATTTATTTATTTAATTAA	4080
Oy	2001	ttcccatgctctaanaatactttttctatctagtataaacatacattgaaacat	2060
Dd	40807	TATATTAATAATTAATATATACATTTATTTATTTAATTAATTTAATTTAT	4086
Oy	2061	aatagttgggaatlcasaatctccatnaatcttttgaactccaaacttata	2120
Dd	40865	TATATTTAATTAATTTATATATATATTAATAATTTAATTAATTAATTAAT	4092
Oy	2121	tttagtcaatacaatgcatalagaagaatcccuaaaaaatttgcatacagaactcca	2180
Dd	40925	TTTAATTAATTTAATATATATATTAATACTAATTAATTAATTAATTTATTTAT	4098
Oy	2181	attcttttttttctatggacaagaalaacacagatcgaaactatttcttctggaa	2240
Dd	40985	ATAATATATATAGTTAATTTATATATTTATTTATTTAATAATTTATTTATGTA	4104
Oy	2241	agttagtaatatacatbaagaanaatttaaaaataatacabaagccatacggccaagt	2300
Dd	41045	TTTAATTAATTTTAGAANAATTTATTTAATTTATATTTTATTAATTTAATTTATTT	4110
Oy	2301	atgtatacttgtagtgytlaataataatgcacgtygcattcagaattgggacaacatg	2360
Dd	41105	ATTATATATATATATATACACTTATATTAATAATTAATTAATTAATTTATTA-TATG	4116
Oy	2361	aaaaacggaattaaaatacttaactttaaaabaaataaaaattggelta	2407
Dd	41164	TATATTAATTAATTTATGAANAATTTATTTAATTTATGTATTTTAAATA	41210

RESULT 11				
PMAL3P5/c				
LOCUS	PMAL3P5	86827 bp	DNA	INV
DEFINITION	Plasmodium falciparum MAL3P5, complete sequence.			
ACCESSION	AL034556	AL008971	AL008978	AL010153
VERSION	AL010206	AL102310	AL139179	AL010162
KEYWORDS	AL034556.3	GI:7711064		
SOURCE	HIG; centromere; CTRP protein; initiation factor E4; Serine/threonine protein phosphatase.			
ORGANISM	malaria parasite P. falciparum.			
REFERENCE	Plasmodium falciparum			
AUTHORS	Eukaryotes: Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
	1 (bases 1 to 86827).			
	Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Fellwell, T., Gierlis, T., Hamilton, R., Hamlin, N., Harris, D., Heyrold, S., Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kyes, S., McLean, J., Moule, S., Mungall, K., Murphy, L., Oliver, K., Quail, M.A., Rajandream, M.-A., Rutter, S., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Whitehead, S., Woodward, J.R., Newbold, C. and Barrett, B.G.			
TITLE	The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum			
JOURNAL	Nature	400	(6744),	532-538 (1999)
MEDLINE	99376085			
REFERENCE	2 (bases 1 to 86827)			
AUTHORS	Bowman, S., Skelton, J., Churcher, C., Lawson, D., Quail, M. and Barrett, B.			

JOURNAL Unpublished
REFERENCE 3 (bases 1 to 86827)
AUTHORS Lawson, D., Bowman, S. and Barrell, B.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1998) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK
COMMENT On May 5, 2000 this sequence version replaced gi:4493931.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum.
Location/Qualifiers
1. 86827
/organism="Plasmodium falciparum"
/strain="3D7"
/db_xref="taxon:5833"
/chromosome="3"
/clone="MAL3P5"
join(322..603,826..1023,1197..1301,1458..2942)
/gene="PFC0575w, MAL3P5.1"
join(322..603,826..1023,1197..1301,1458..2942)
/gene="PFC0575w, MAL3P5.1"
/note="PFC0575w (MAL3P5.1), Hypothetical protein, len: 689
aa, possible signal sequence, revised: added new exon 2"
/codon_start=1
/product="Hypothetical protein, PFC0575w"
/protein_id="CAB38969.2"
/db_xref="GI:7711065"
/db_xref="SPTREMBL:O97258"
/translation="MYLKNVYIYISCFILFDLCFSFHLKMKYKNNMKNKSVTEFL
RSPOIYRKRKRSRIRKNVSPFKOKKPLFENLKQCSFGWRNOYDQYIDVIS
NINNLPIKQVTHKKKNEFTKEIKQILHCVSKIDFKIINLSYIIKHQMSNT
VHSLNIOSEKVEKDAENYLAHLFLKIDENTLPSMHIIMDFEKKQVIECID
IKSKKKRKLSTIYNLFICTLYCTQCMCLIKIYSHLCLIFEFEEFPCFLCNILIER
IYECVGDILRKIRERYNLICEKKIKTFHMKADAKKEITMKDDIYFNHYHDLILC
FTRKLIERNKNNKNIIRSNYNDININDISIDKDMNNTIDVINNISIDEKIKDEEFP
DDENLKLDYEQFQLENDNIIKYIEEDPLXNINNSNINNNININIKKHKIK
DYNDDDDYEEKEDYIQKNDIDYIRKNTIGKNNKLEEFKQFIDQADIEFQNS
NVNLDQGRVKSNDENKSTEHKNTINKGDTLQIQOMENEFIRKNDINISND
ISNNEINIIRKLKLNOSDEDINLSDIYERLRKVLWYIOKIEYKFOYDITINDQ
YPIIKNKTYLDLNTGYKTYMSPDVNSLFEKTKIDISINEKDKNNQMNQKSKNY
TIMPNVW"
604..609
/note="predicted splice donor sequence for exon 1 of
PFC0575w (revised)"
816..825
/note="predicted splice acceptor sequence for exon 2
(revised of PFC0575w)"
1024..1029
/note="predicted splice donor sequence for exon 2 of
PFC0575w (revised)"
complement(3354..6644)
/gene="PFC0580c, MAL3P5.2"
complement(3354..6644)
/gene="PFC0580c, MAL3P5.2"
/note="PFC0580c (MAL3P5.2), Hypothetical protein, len:
1097 aa, possible signal sequence, predicted using
hexExon"
/codon_start=1
/product="Hypothetical protein, PFC0580c"
/protein_id="CAB38968.1"
/db_xref="GI:4493932"
/db_xref="SPTREMBL:O97257"
/translation="MLGILKRNVLIVSVSLFAVFLKHKNNENYELLINNDIE
KIKIRIHNKCSIPLPLFNIDSYIRKNNILRLYKFRKRRDKDEYYITIMVK
KRRRAIKYNTISDEQNLFNKFYIEIVLEYSLKGLISPHLSYILKNSHCVNISC
SLYYNKLONKHNILNEKRLYKFOINNEHQAPGTHHTNNNNNNKPKPIDIINHC
KNTNISSTYNNNEKENINIDKYNIHNFTKESISYKDCQHITILMYLILNQY
DNICRIQNTNININENYMINILKYCYKMETIILNYNHIEQMKKINOKNNTNS
LFKYISFPEFKENHIDYLFEDQNMNHLKREKNDKFPYNSNNTNNIKYISDNY
FYHINSSNRCSFKNLKQOOTDNTKHTIIGKSKRYMNSDSHEKNNNTGCGNINIK
DQKDLILKTYFLKGNLDDIQLINLEYIMYMLLECSLISIKKNIHLLEKKE
FDKNKIITYNSADYNNNLRNLIKRSKNEERENISFASPLLSKNIIYFEDIEG
RSRDNITYNVYDKETNKTNNNNNDNDNICSNNDHICSNNDHICSNNDHICSN

NNNNICSNNNNNICSNNNNNICSNNNNNICSNKMLDFFCODKNFNDYNTKKKKRRI
YELAKITYNNIPFLYKGGKKEKHONEDNTIMLYIKKKEPMIPIYKLIIRKDDSPIH
NNNIVNDGIDTNNNIIPEKKRYNLESSISISFYIKDIEYKRYKRLXYINLKKRKR
YFLKMNHIRKRLKMKRATHTYITNFDQITINNYITIKKNIHKHILHKKQCKRE
IDIVKFDILYCCMINNININIFSYIHKDHNHCYRIKRAVKNLILYEYNLKNENY
YKNNKIKRYLTYNNNNISNDLYPNISYKNIYONKRYNRNFSHIESLNDHLILF
LYTORIONCDDISFYIKRYNNEKNFLYLYELHYIYSEKKKKKKPFSSSD
YSYDTONSFTFSYFFFSYSLYFLFHYHDMVASYIFFTLYSGLPYYSLYNNI
MYVCGPRT"
join(7669..7717,7800..7829,7912..7940,8064..8106,
8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,
9369..9505,9613..9838)
/gene="PFC0581w"
join(7669..7717,7800..7829,7912..7940,8064..8106,
8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,
9369..9505,9613..9838)
/gene="PFC0581w"
/note="PFC0581w, hypothetical protein, len: 324 aa,
revised: new gene prediction, splicing prediction very
tentative"
/codon_start=1
/product="hypothetical protein, PFC0581w"
/protein_id="CAB90285.1"
/db_xref="GI:7711066"
/translation="MKKKKHYSIKFVNDKRYKYSKDYIRIINWIGLHIFRL
OKIFHYEANDQIASFFESHKNRYRVRKEDIQIECKWNIDYLSIESIKODNS
NDIKTYTAKNNKNDQLLTTSYKRIIDTYPNASQWQSEYIFLNKRFHSRNSPRA
LKXKDEIVSKNNFSALSNDNSYKTYIDYLDLNDIIESEYKYNASGVKVV
TLKREKRIWNRLLSEKYPNMQVMDMEKRIHTIYPTVNLIFLISLSHRHDSV
QNFLEKEKNSDKLDQIDDEDEKYPDEILREAKKSEEDKDEEL"
7790..7799
/note="potential splice acceptor sequence for exon 2 of
PFC0581w"
7830..7835
/note="potential splice donor sequence for exon 2 of
PFC0581w"
7901..7911
/note="potential splice acceptor sequence for exon 3 of
PFC0581w"
7941..7947
/note="potential splice donor sequence for exon 3 of
PFC0581w"
8055..8063
/note="potential splice acceptor sequence for exon 4 of
PFC0581w"
8107..8112
/note="potential splice donor sequence for exon 4 of
PFC0581w"
8293..8301
/note="potential splice acceptor sequence for exon 5 of
PFC0581w"
8369..8374
/note="potential splice donor sequence for exon 5 of
PFC0581w"
8559..8567
/note="potential splice acceptor sequence for exon 6 of
PFC0581w"
8642..8647
/note="potential splice donor sequence for exon 6 of
PFC0581w"
8704..8712
/note="potential splice acceptor sequence for exon 7 of
PFC0581w"
8813..8818
/note="potential splice donor sequence for exon 7 of
PFC0581w"
8913..8923
/note="potential splice acceptor sequence for exon 8 of
PFC0581w"
9007..9012
/note="potential splice donor sequence for exon 8 of
PFC0581w"
9112..9121
misc_feature

/note="potential splice acceptor sequence for exon 9 of PFC0581w"
9259..9264
/note="potential splice donor sequence for exon 9 of PFC0581w"
9361..9368
misc-feature
/note="potential splice acceptor sequence for exon 10 of PFC0581w"
9506..9511
/note="potential splice donor sequence for exon 10 of PFC0581w"
9601..9612
misc-feature
/note="potential splice acceptor sequence for exon 11 of PFC0581w"
complement(join(10024..10044,10320..10336,10397..10445,10613..10674,10773..10798,10929..10999,11157..11237,11453..11515,11715..11767,11950..12040))
/gene="PFC0582c"
complement(join(10024..10044,10320..10336,10397..10445,10613..10674,10773..10798,10929..10999,11157..11237,11453..11515,11715..11767,11950..12040))
/gene="PFC0582c"
/note="PFC0582c, hypothetical protein, len: 177 aa, revised: new gene prediction"
/codon_start=-1
/product="hypothetical protein, PFC0582c"
/protein_id="CAB90286.1"
/db_xref="GI:7711067"
Query Match 1.8%; Score 125.4; DB 96; Length 86827;
Best Local Similarity 44.0%; Pred. No. 2.5e-08;
Matches 1033; Conservative 0; Mismatches 1266; Indels 38; Gaps 10;
QY 108 ataataaataaataatcattcaataataatcagaagtaataatcattcattcgtt 167
DB 41351 ATAAATTAATTTGCAATTTATTAATAATTAATTTATTTCAAAATTTATTAATTA 41292
QY 168 atgctataaacaataatcattccttggtatagagaataatgaggttggaattata 227
DB 41291 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 41232
QY 228 ataataaaggaataatcagctcattcgttggtatcagcttgaagtttggttc 287
DB 41231 ATTTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTAT 41172
QY 288 ttctgtatatagtatatacgagtaatacaaaagagatcgttggaagtgtaacatactc 347
DB 41171 TAAATATACATTAATAATAATTAATAATTAATAATTAATAATTAATAATTAATA 41121
QY 348 gttatgaccccccaaaaaaacaacaaacaaaccccccccccgatatagttt 407
DB 41120 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 41061
QY 408 ttggttcgtagtaggtttcattcgtacataatcagcattccttggttactatg 467
DB 41060 TCATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 41001
QY 468 aagaattcctaccataataatcgaatcattcctccttgctt---ataataataat 524
DB 41000 TAACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40941
QY 525 acgagtgtagatccgtttacgacacccacatcagtgtagtgcttggtgcatac 584
DB 40940 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40881
QY 585 cagcaaatcattcaacaagagatcgtgaaacaaacgaaataaagaaaggaagtag 644
DB 40880 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 40821
QY 645 tgaaccatgagtagtctgaaatcattcaagaagaataagagatgcacaaaggttg 704
DB 40820 TAA--ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40763

QY 705 tggataaagtgccctgcagccttctcacaataatcagaccatttgatcttc 764
DB 40762 TTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40703
QY 765 tggatattcgtaaaattcgcgaatcagatcgtgaaataattcttgtagccgac 824
DB 40702 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40643
QY 825 tcaattatgctccaggtatcgtcacaatcctcgttttaagcattcttgcttcct 884
DB 40642 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 40583
QY 885 ttggttcgttctcctaataatataatcgcgataatgataaagatataatcac 944
DB 40582 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40523
QY 945 aaacacattgctcggagccatttgaaataactttcccaacattcagggacactg 1004
DB 40522 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40463
QY 1005 ctgcaccccttaaaatacgaatttaccagcgtcactagctgagattcagcataagcata 1064
DB 40462 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40403
QY 1065 aagagaccgttcgaactattatacaagtltaacaacatgaatagcttgaaacctta 1124
DB 40402 AATTAAGTTTAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40343
QY 1125 gaaatttggaaattccggttgcttgtaataatagatttagtga--aacaatattg 1183
DB 40342 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40283
QY 1184 taatcaattagtgccaacataacataatccttccacaagaaacaaactaagaag 1243
DB 40282 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 40223
QY 1244 ttaacatccataatataatggtatgatactacattcagctatgatacagagaccag 1303
DB 40222 AGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40165
QY 1304 aatagatagtgatgcgttaataatgaattcacaacgctggttaataatcagagccgta 1363
DB 40165 ----GTATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40111
QY 1364 tgtacgactcgcgaataatcattccttggttggtcacaataaacaacaaagaa 1423
DB 40110 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 40051
QY 1424 aagaagaacgatttctcgttgatcattcaatgactcaaaatgcatagatccttggtg 1483
DB 40050 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39991
QY 1484 ttaacgttcgcgaagctccctacacagcgtgtaacacatcgaactataatcttctt 1543
DB 39990 TATTAAGTAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39931
QY 1544 taatgcattcctaacaattatttggttggaattcaataaagcggaactgtaacat 1603
DB 39930 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39872
QY 1604 taacaattatataatgatactagatcgtgatactccaatcacaactcgttgatgta 1663
DB 39871 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39812
QY 1664 aacttaactgttctcctcagataataatataatcagaggttaaaaaagtttg 1723
DB 39811 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39760
QY 1724 tcttatttcgcgagtcgaagaagataaacctaaatgactttaatttcttgaaatgtaac 1783
DB 39759 GCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39700
QY 1784 cctttaccatagatataatcagatgcttcttggtgcataatgacagccctcacaac 1843

REFERENCE
AUTHORS
TITLE
2 (bases 5269 to 5695)
Clary,D.O., Wahleithner,J.A. and Wolstenholme,D.R.
Transfer RNA genes in *Drosophila* mitochondrial DNA: related 5' flanking sequences and comparisons to mammalian mitochondrial tRNA genes
Nucleic Acids Res. 11 (8), 2411-2425 (1983)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
3 (bases 404 to 5272)
de Bruijn,M.H.
Drosophila melanogaster mitochondrial DNA, a novel organization and genetic code
Nature 304 (5923), 234-241 (1983)

JOURNAL
MEDLINE
AUTHORS
TITLE
4 (bases 804 to 1778)
Satta,Y., Ishiwa,H. and Chigusa,S.I.
Analysis of nucleotide substitutions of mitochondrial DNAs in *Drosophila* melanogaster and its sibling species
Mol. Biol. Evol. 4 (6), 638-650 (1987)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
5 (bases 5268 to 13619)
Garesse,R.
Drosophila melanogaster mitochondrial DNA: gene organization and evolutionary considerations
Genetics 118 (4), 649-663 (1988)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
6 (bases 441 to 2967)
Satta,Y. and Takahata,N.
Evolution of *Drosophila* mitochondrial DNA and the history of the melanogaster subgroup
Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9558-9562 (1990)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
7 (bases 14215 to 14512)
Ballard,J.W., Olsen,G.J., Falth,D.P., Odgers,W.A., Rowell,D.M. and Atkinson,P.W.
Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods
Science 258 (5086), 1345-1348 (1992)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
8 (bases 14917 to 19517)
Lewis,D.L., Farr,C.L., Farguhar,A.L. and Kaguni,L.S.
Sequence, organization, and evolution of the A+T region of *Drosophila* melanogaster mitochondrial DNA
Mol. Biol. Evol. 11 (3), 523-538 (1994)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
9 (bases 1 to 408; 13319 to 19517)
Lewis,D.L., Farr,C.L. and Kaguni,L.S.
Drosophila melanogaster mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons
Insect Mol. Biol. 4 (4), 263-278 (1995)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
10 (bases 1 to 19517)
Lewis,D.L., Farr,C.L. and Kaguni,L.S.
Direct Submission
Submitted (03-Oct-1995) Laurie S. Kaguni, Biochemistry Department, Michigan State University, East Lansing, MI 48824-1319, USA
Location/Qualifiers

FEATURES
source
1. 19517
/organism="Drosophila melanogaster"
/organelle="mitochondrion"
/db_xref="taxon:7227"
/note="derived from new and previously submitted sequences: sequence is a composite containing sequences obtained from different *Drosophila* melanogaster strains"

1. 65
/product="tRNA-Ile"
complement(97..165)
/product="tRNA-Gln"
171..239
/product="tRNA-Phe"
240..1265
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 2"

CDS

/protein_id="AAC47811.1"
/db_xref="GI:1166530"
/translation="MPNNSSKILFITTNICTLTYSNSGLAMGLEINLSFIP
LSDNNMLSTESLKYFLVYLASTVILFSSITLMLKNNKNNNEISFSTINLSALL
LKSGAAPFHEWEPMMMEGLTMMNALMTWKIAPLMLISLNTKYLILLSVLISVIT
GAIGLNTLSRLKLMARFSSINHLGMLSSIMSSIMLILEFFSFLSFLTEFWNF
KLPHNOLSEWFSNKSILKFTFLPNNPLSLGLPELPKPLVIOOLTCONOYFML
IMMSLTITLLEFVIRICYSAPFMNPFENNIMKNNMSININMIMTFPSIRPLFLI
SLFFPM"

1264..1329
/product="tRNA-Trp"
complement(1322..1383)
/product="tRNA-Cys"
complement(1403..1468)
/product="tRNA-Tyr"
join(1470..1472,1474..3009)
/note="mechanism underlying reading frame shift after first codon uncertain"

/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit I"
/protein_id="AAC47812.2"
/db_xref="GI:7412849"
/translation="MSROWLSTNHNKDIGLYFFGAMAGVGTSLILTRAEIGHPC
ALIGDDIYNYTVTAHAFIMIEFMVPMIGFGNMVPLMLGAPDMAFPMNMSFW
LLPALSILLVSSWVNGACGVTVPYPLSAGINAGSADIATFSLHAGISITLGA
VNFITTYNNRSTGISLDRPLFVSWVITLALLSLIPYAGATMLLDRNLNTSF
FDPAGGDDPLLYQLHWFEGHPEVYILILPEFGMISHISOESGKRTFSLGMITVM
LAIDLGEIVAHMFTVGMDDVDRAYTSATMIJAVTGKIFSMATLHGDTLSYS
PAIIMALGEVFLFTVGGLGVLANSDIILHDTYVVAHFHVLMSGAVFAIMAF
IHWPLFTGLTKNKKMLKSHFTIMEIGVNLFPFOHFLGLAGMPRYSDDVATYTN
IVSTIGSTISLGLTFPFFITWESLVQROQYIYIQLNSTIEWQNTPRAHSSELP
LLTN"

3012..3077
/product="tRNA-Leu"
3083..3767
/note="TAA stop codon is completed by the addition of 3' A residues to the mRNA"
/codon_start=1
/transl_except="(pos:3767,aa:TERM)
/transl_table=5
/product="cytochrome c oxidase subunit II"
/protein_id="AAC47813.1"
/db_xref="GI:1166532"
/translation="MSTWANGLDSDASPLMEQLIFPHDALLILWATYVGLMER
LEFNNYVNRFLHGLIEMITLIPAILLFTALPSRLLLDEINEPSVTLKSIGH
QWVSYSDENNTIEFDSYMTPLNELMTDGERLDVNRVYLPNNSOIRILVTADVI
HSWVPALGVVDTGTPRLNQTFFINRPGILFYQCSEICGANSFMPYIESVPVY
FIKWSNNS"
3768..3838
/product="tRNA-Lys"
3840..3906
/product="tRNA-Asp"
3907..4068
/codon_start=1
/transl_table=5
/product="tRNA-Phe"
/protein_id="AAC47814.1"
/db_xref="GI:1166533"
/translation="MPQMAPISWLLPFIISITPLRCSINYSYMPNPSPKNELKNI
NLNSMNKMW"
4062..4736
/codon_start=1
/transl_table=5
/product="tRNA-Pase 6"
/protein_id="AAC47815.1"
/db_xref="GI:1166534"
/translation="MMTNLPSVDPDLAIFNLSLWLSPLGLIMIPSYIWLMPRSYNI
MMNSILFLHKEFTLILGSPCHNSSTFIFISFLILFNNMGFLPIFTSTSLITLT
LSLADPLMLCPKLTGINTNTHOHAHLVPGSTPAIIMPFCITFTSNTIPGLTAR
LTANNIAGHLTLTLTGNTSSSYMLTFLMAOIALLVLESAAVAMIOISYVAVLSPL
YSSFVN"

4736..5524

CDS

1607 aatattatattagatagatgtaagttatctccaataacactacttgtagttaaac 1666
18737 AATTAAAAAGTAGATATTAATTTAAAAATTTATATTCATATTTATTTATTTAA 18796
1667 ttaactctgtctctcctacggtataaalattacatcagsggtaaaaaagttgct 1726
18797 TTTAATTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 18856
1727 tatttcgcgtagcagagataaactaatgacttaatttttgaataatgtaccc 1786
18857 ATTATGGAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 18915
1787 cttaactagatgaataactgactgtcttgctccataagagactcacaacgt 1846
18916 TTTAATTAATTAATTTTAAAAATTTCTTAATGTAATTTTAAATAATTAATTAAT 18975
1847 gatagtcaattttctgcgaataataattaggaattcaatgcactacataaga 1906
18976 AATAAATCATTTTATTTAAAAATTAACAAAAATTTTAAATAATTAATTAATG 19035
1907 aacgctgagatcattcaatttaagaacaaaatttggaaaaatgttaattct 1966
19036 AAATTAATTTATTAATTTTCAATTTTAAAAAAATTTTAAAAAAATTAATTTT 19095
1967 aacaatattataaataatgactgaatgactatgctcctatgcttaaaatatt 2026
19096 TTTTAAAAAAACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 19135
2027 ttatattagttataaataacattatgaacataatggttggaattcaatactcc 2086
19156 TTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19215
2087 attaatatttttgaataactcaaatattatattgactgaataacatgctaga 2146
19216 ATATTAATTTAAAAATTTAATTTATTTATTTATTTAATTTAATTTAATTA 19275
2147 ttccaataaatttgttgaagaacttccaatttttttttttttgaagaaga 2206
19276 TAAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19335
2207 taacagatagaataactatttgttggaagtagtaataacatttaagaacatt 2266
19336 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 19395
2267 taaaaaatatataagcctatacgccgctcaaaagtgtatctagtgagtgtaata 2326
19396 AAAAAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAA 19455
2327 atgcagtgtgcattcgaattggagcaacatgaacaggaattaaatatlaact 2386
19456 TTTTAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19515

RESULT 14
PFMALIP3/c
LOCUS PFMALIP3 67970 bp DNA INY 15-DEC-1999
DEFINITION Plasmodium falciparum MALIP3, complete sequence.
ACCESSION AL031746
VERSION AL031746.9 GI:6594243
KEYWORDS HTG.
SOURCE malaria parasite P. falciparum.
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 67970)
AUTHORS Bowman, S., Churcher, C., Harris, B., Harris, D., Lawson, D., Quail, M., and Barrell, B.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-1998) P. falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
COMMENT On Dec 16, 1999 this sequence version replaced gi:5763807.
For more information about this sequence or the Malaria Project,

see http://www.sanger.ac.uk/Projects/P_falciparum. IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E. coli, yeast, vector, phage etc.

FEATURES
Source

1. 67970
/organism="Plasmodium falciparum"
/strain="3D7"
/db_xref="taxon:5833"
/chromosome="1"
complement(join(1748. 2598, 2748. 2848, 2990. 3276))
/gene="MALIP3.01"
complement(join(1748. 2598, 2748. 2848, 2990. 3276))
/note="MALIP3.01, conserved hypothetical protein, len: 412 aa, similarity: UPF0006 family eg to YB1055C/YB10512/YB10511, YB85, YEA5T (418 aa), fasta scores: opt: 316, E(): 1.1e-12, (33.28 identity in 271 aa overlap)"
/codon_start=1
/product="conserved hypothetical protein, UPF0006 family"
/protein_id="CAB63556.1"
/db_xref="GI:6594244"
/translation="MKLVFHYIKYINULFIYSIFLKSNSKITNDLRYISTVKKYV LQIKRSLMKNNHNRKMEDESSFDIGSWLTKMEGVNSKKHEDLONVNRK NNNVDKIIITCTCLAEIDKSLKICEYDPEKFLYLSAGVPTNCYEDIDNKHEKE IIAKREVEEFKYKNEGVNSKMGNGKRCDCDEKMMNNNEILLENIDTIPGKY NEKREYVENLKNKIKYIPNIVICIGEGFDLSEKSKYIOIKYFPOKLQVFN LPMFLHMGNCSEFPKIKYIDYKFLPKRNGYIHSEFTPKEDIVHIVONKLYGVNG CSLSLENINAVKKIPLNLLLELDAPGCGVKKTHASIEYIKDYEKATYNLKKIK NLIKDDNTIFERNEPYNIA
complement(2599. 2610)
/note="potential splice acceptor sequence"
complement(2742. 2747)
/note="potential splice donor sequence, atg/gttaa"
complement(2849. 2861)
/note="potential splice acceptor sequence"
complement(2984. 2989)
/note="potential splice donor sequence, aaa/gtaaaa"
5005. 5496
/gene="MALIP3.02"
5005. 5496
/gene="MALIP3.02, hypothetical protein, len: 163 aa, contains possible signal sequence"
/codon_start=1
/product="hypothetical protein, MALIP3.02"
/protein_id="CAB63557.1"
/db_xref="GI:6594245"
/translation="WKLNNRFVVLCPILIFFFLNSVGLGNNNNRNINFEHENA KAMRLLSGEINSIKLDNGDELKIKLNDEKHDSYKWSYSEFISNLEBEKYSQD LFKKORINEANTKIIDROEFYIINDEIENIATRYVLENNPFLYIQSPKQSLID IGS LNN"
8020. 10389
/note="possible cen1, region of very high [A+T] content"
14884. 20352
/gene="MALIP3.03"
14884. 20352
/gene="MALIP3.03"
/note="MALIP3.03, putative ABC transporter, len: 1822 aa"
/codon_start=1
/product="putative ABC transporter"
/protein_id="CAB63558.1"
/db_xref="GI:6594246"
/translation="MTYKKENVGTSNGKNGKKKSCNIFLNFSPMIRPLINDLIK GDIDELPNICNPVPYVYASKLEENLRDIEVDEDFYSEKSSNENHVIHRNSDAS EKKVNYIHNLWLSILTKFKRILILISFYLITLVLTGKFLDYFWRILIEGKIPV YISFLKDFKFSGLVVMVIMFHLFFALFLAFYHFLTINLKVIMFLVKINCSNN NHIQNDPAFYNTYRRFSQPTIDEISDFLSIGNASSSSSGICNNKNDNNFEVEN DYIINFKSTKMKKDSLNENRSLPNVNIYIMSDVPSVYFVFTSCINLENVVKLF

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2001, 18:43:30 : Search time 27.59 Seconds
(without alignments)
2460.039 Million cell updates/sec

Title: US-09-502-426-2
Perfect score: 2681
Sequence: 1 MFETEHHTLLPLLLPSLTS.....FAFPVDFPNGLRVSRLL 513

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_16:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2681	100.0	513	10	064989 arabidopsis
2	2677	99.9	513	10	09SC09 arabidopsis
3	951	35.5	474	10	09LKH7 phaseolus a
4	866	32.3	512	10	09FX29 arabidopsis
5	864	32.2	524	10	023242 arabidopsis
6	860	32.1	457	10	09M066 arabidopsis
7	822	30.7	478	10	09LW73 arabidopsis
8	815	30.4	482	10	09LY89 arabidopsis
9	789	29.4	464	10	09LIC5 arabidopsis
10	726	27.1	465	10	09FMA5 arabidopsis
11	711	26.5	465	10	09LH81 arabidopsis
12	658	24.5	463	10	09FH76 arabidopsis
13	632	23.6	457	10	065624 arabidopsis
14	631	23.5	518	10	09FI38 arabidopsis
15	631	23.5	735	10	09LGI7 arabidopsis
16	622	23.2	443	10	09LJK2 arabidopsis
17	611	22.8	482	10	081077 arabidopsis
18	589	22.0	486	10	09LIF9 arabidopsis
19	581	21.7	487	10	023384 arabidopsis

20	581	21.7	496	10	09FOY4 arabidopsis
21	579.5	21.6	474	10	09SHY7 arabidopsis
22	579.5	21.6	485	10	09SJH2 arabidopsis
23	577	21.5	455	10	09LXH8 arabidopsis
24	575.5	21.5	464	10	004949 arabidopsis
25	569.5	21.2	477	10	09LYV7 arabidopsis
26	547.5	20.4	489	10	09ZV72 arabidopsis
27	518	19.3	504	10	09SNG3 oryza sativ
28	497.5	18.6	460	10	09SYN2 arabidopsis
29	479.5	17.9	444	2	09SYN2 arabidopsis
30	461	17.2	497	11	09RIE4 oryza sativ
31	446	16.6	512	4	09NR63 xenopus lae
32	444.5	16.6	492	13	093323 xenopus lae
33	442.5	16.5	525	4	09NPA1 xenopus lae
34	439.5	16.4	492	13	09PUB4 xenopus lae
35	435	16.2	224	10	09SDM6 xenopus lae
36	398.5	14.9	375	10	09LW32 arabidopsis
37	369	13.8	349	10	09FWR7 arabidopsis
38	357	13.3	518	5	017807 caenorhabdi
39	338	12.6	453	2	09KFA6 caenorhabdi
40	329.5	12.3	517	5	017806 caenorhabdi
41	321	12.0	497	13	09PVI0 oryza sativ
42	317	11.8	483	10	09FW33 oryza sativ
43	305	11.4	501	11	09QXF7 ratu mus
44	300	11.2	491	11	064463 mus muscu
45	300	11.2	513	10	09SBP8 populus tri

ALIGNMENTS

RESULT 1	
064989	PRELIMINARY: PRT: 513 AA.
AC 064989	
DT 01-AUG-1998 (TREMBLrel. 07, Created)	
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)	
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
DE STEROID 22-ALPHA-HYDROXYLASE.	
GN DMF4.	
OS Arabidopsis thaliana (Mouse-ear cress).	
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;	
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;	
OC Brassicales; Brassicaceae; Arabidopsis.	
OX NCBI_TaxID=3702;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=WS-2;	
RA Choe S., Dikes B.P., Fujioka S., Takatsuto S., Sakurai A.,	
RA Feldmann K.A.;	
RL Plant Cell 10:0-0(1998).	
CC -1. SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.	
DR EMBL; AF044216; AAC05093.1; -.	
DR InterPro: IPR001128; -.	
DR Pfam: PF00067; P450; 1.	
DR PROSITE: PS00086; CYTOCHROME P450; UNKNOWN_1.	
KW Heme; Monooxygenase; Oxidoreductase.	
SQ SEQUENCE 513 AA: 58667 MW; B1639BDD9A4DA6F3 CRC64;	
Query Match	100.0%; Score 2681; DB 10; Length 513;
Best Local Similarity	100.0%; Pred. No. 6.2e-183;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY 1 MFETEHHTLLPLLLPSLTSLLPLLLKRRKRKRFNLPFGKSGMPFGGTGTYKPT 60	09FGY4 cucurbita m
	09SHY7 arabidopsis
	09SJH2 arabidopsis
	09LXH8 arabidopsis
	004949 arabidopsis
	09LYV7 arabidopsis
	09ZV72 arabidopsis
	09SNG3 oryza sativ
	09SYN2 arabidopsis
	09SYN2 arabidopsis
	09RIE4 oryza sativ
	09NR63 xenopus lae
	093323 xenopus lae
	09NPA1 xenopus lae
	09PUB4 xenopus lae
	09SDM6 xenopus lae
	09LW32 arabidopsis
	09FWR7 arabidopsis
	017807 caenorhabdi
	09KFA6 caenorhabdi
	017806 caenorhabdi
	09PVI0 oryza sativ
	09FW33 oryza sativ
	09QXF7 ratu mus
	064463 mus muscu
	09SBP8 populus tri
OY 61 ATTLGDFMOOHVSKYKGYKIRSNLFCGPTIVSADAGLNRRITLNEGRIFPCSPRSIGTL 120	
OY 61 ATTLGDFMOOHVSKYKGYKIRSNLFCGPTIVSADAGLNRRITLNEGRIFPCSPRSIGTL 120	
OY 121 GRWSMLVYLVDGMDHRMRSISLNLFLSHARLRTLLKDVRRHTLFLVDSWOONSIFSAODEA 180	

```

Db 121 GKMSMLVVGDMHRDMSISLNFLSHARLTKILKDVERTLFLVDSMOONSIFSADDEA 180
QY 181 KKFENFLMAKHMSMDGEEETBOLKKEYVTFMKGVSAFNLPGTAVHKAQSRATILK 240
Db 181 KKFENFLMAKHMSMDGEEETBOLKKEYVTFMKGVSAFNLPGTAVHKAQSRATILK 240
QY 241 FIEKMEERKLDIKEEDQEEBEVTEDEAEKSDHVRKQRTDDLLGWLKHNSLSTEQ 300
Db 241 FIEKMEERKLDIKEEDQEEBEVTEDEAEKSDHVRKQRTDDLLGWLKHNSLSTEQ 300
QY 301 ILDLISLFLAGHETSSVAIALAIFLQACPKAVEELREHLELTARAKKEGSELNMD 360
Db 301 ILDLISLFLAGHETSSVAIALAIFLQACPKAVEELREHLELTARAKKEGSELNMD 360
QY 361 YKKMDFQVCYINETLRIGNVYRFLHRKALDVRKGYDIPSGMKVLPVISVHLNDSRYD 420
Db 361 YKKMDFQVCYINETLRIGNVYRFLHRKALDVRKGYDIPSGMKVLPVISVHLNDSRYD 420
QY 421 QPNLFNPMWQOONNGASSGSGSFTWGNMYPFGGPRLCAGSELAKLEMAVFIHHLV 480
Db 421 QPNLFNPMWQOONNGASSGSGSFTWGNMYPFGGPRLCAGSELAKLEMAVFIHHLV 480
QY 481 LKENWELAEDDQPPAPFPVDPNGLPIRVSRIL 513
Db 481 LKENWELAEDDQPPAPFPVDPNGLPIRVSRIL 513

```

RESULT 2

```

09SC09 PRELIMINARY; PRT: 513 AA.

```

```

AC 09SC09;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE STEROID 22-ALPHA-HYDROXYLASE (DMF4).
GN T3A5.40.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bioecker H., Mewes H.W., Lemcke K., Mayer K.F.X., Quetler F.,
RA Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC EMBL: AL132979; CAB62435.1; -
DR InterPro: IPR001128; -
DR Pfam: PF00067; P450.1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 513 AA; 58867 MW; B1639BD9A5D7C93 CRC64;

```

```

Query Match          99.9%; Score 2677; DB 10; Length 513;
Best Local Similarity 99.8%; Pred. No. 1.2e-182;
Matches 512; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MFEETHHTLPLLLPLSLLSLLFLILKRRNRKTRFNLPKSGMFPGLGTIGLYLPPY 60
Db 1 MFEETHHTLPLLLPLSLLSLLFLILKRRNRKTRFNLPKSGMFPGLGTIGLYLPPY 60
QY 61 ATTLDGDMQOHVSKYKGRYSNLEGEPTIVSADAGLNRFILQNGRLFECSYPSRIGIL 120
Db 61 ATTLDGDMQOHVSKYKGRYSNLEGEPTIVSADAGLNRFILQNGRLFECSYPSRIGIL 120

```

```

QY 121 GKMSMLVVGDMHRDMSISLNFLSHARLTKILKDVERTLFLVDSMOONSIFSADDEA 180
Db 121 GKMSMLVVGDMHRDMSISLNFLSHARLTKILKDVERTLFLVDSMOONSIFSADDEA 180
QY 181 KKFENFLMAKHMSMDGEEETBOLKKEYVTFMKGVSAFNLPGTAVHKAQSRATILK 240
Db 181 KKFENFLMAKHMSMDGEEETBOLKKEYVTFMKGVSAFNLPGTAVHKAQSRATILK 240
QY 241 FIEKMEERKLDIKEEDQEEBEVTEDEAEKSDHVRKQRTDDLLGWLKHNSLSTEQ 300
Db 241 FIEKMEERKLDIKEEDQEEBEVTEDEAEKSDHVRKQRTDDLLGWLKHNSLSTEQ 300
QY 301 ILDLISLFLAGHETSSVAIALAIFLQACPKAVEELREHLELTARAKKEGSELNMD 360
Db 301 ILDLISLFLAGHETSSVAIALAIFLQACPKAVEELREHLELTARAKKEGSELNMD 360
QY 361 YKKMDFQVCYINETLRIGNVYRFLHRKALDVRKGYDIPSGMKVLPVISVHLNDSRYD 420
Db 361 YKKMDFQVCYINETLRIGNVYRFLHRKALDVRKGYDIPSGMKVLPVISVHLNDSRYD 420
QY 421 QPNLFNPMWQOONNGASSGSGSFTWGNMYPFGGPRLCAGSELAKLEMAVFIHHLV 480
Db 421 QPNLFNPMWQOONNGASSGSGSFTWGNMYPFGGPRLCAGSELAKLEMAVFIHHLV 480
QY 481 LKENWELAEDDQPPAPFPVDPNGLPIRVSRIL 513
Db 481 LKENWELAEDDQPPAPFPVDPNGLPIRVSRIL 513

```

RESULT 3

```

09LKH7 PRELIMINARY; PRT: 474 AA.

```

```

AC 09LKH7;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CYTOCHROME P450.
GN CYP7.
OS Phaseolus aureus (Mung bean) (Vigna radiata).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Vigna.
OX NCBI_TaxID=3916;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang M.T., Chen Y.M.;
RL Cloning and sequencing of a Vigna radiata cDNA encoding cytochrome
RT P450.
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC OXIDIZED FLAVOPROTEIN + H(2)O (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM (BY
CC SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: AF279252; AAF89209.1; -
DR InterPro: IPR001128; -
DR Pfam: PF00067; P450.2.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Electron transport; Endoplasmic reticulum; Heme; Membrane; Microsome;
KW Monooxygenase; Oxidoreductase.
SQ SEQUENCE 474 AA; 54037 MW; A2542A809C5BAC6D CRC64;

```

```

Query Match          35.5%; Score 953; DB 10; Length 474;
Best Local Similarity 40.1%; Pred. No. 8.2e-60;
Matches 203; Conservative 97; Mismatches 162; Indels 44; Gaps 10;

```

```

QY 12 LLLPLSLLSLLF-----LILKRRNRKTRFNLPKSGMFPGLGTIGLYLPPYATTGLD 66
Db 1 MVSLEPLT--LLFLAASAAALFLHRAFSRRKFLRPPGSGYGLPFGITQLLSAVKSSNPEP 58

```

```

QY 67 FMOQHSKGIYRSNLGFEPTIVSADAGLNRFILONEGRLEFCSYPRSIGGLGKWSML 126
DB 59 FMDENRRRGSLFPMTHVEBEPTVFSADPELNRFILONEGRKLDSCSPGISNLLGKHSLL 118
QY 127 VLVGMHMRMSISLNLSHARLRTLLKDOVERHTLFLDWSQOONSIFSADAEKAKFTFN 186
DB 119 LKKGALHKRMSHSLTMSFANSIILKDLHHIDRLIGLNDJWSDR--VILMOAKKITPE 176
QY 187 LMAKHMSMDPEEETEOLKKEYVPMKGVNASPLNLPCTAHHKALQSAATILKFERKM 246
DB 177 LVYKQMSMDP--DEWESLRKEVLYEGFPLPLPSTFYRRALKARFKVAEALTIV 235
QY 247 EERKLDIKEDEDEEVEKTEDEAEMSKSDHVRKORTDOLLGWL--KHSNLSSTEQILDI 305
DB 236 RGR---REYNOGKEKKS-----DMGLALLSGHFDSDQIVDL 272
QY 306 LSLLRPHGHTSSVALALAIFFLOACPRAVEELREHLEIARAKKEIGESLMWDYKKKD 365
DB 273 LALLVAGYETSTIMTLAVKFLTEPLPLALQKEHDOJ--RARSDEG--APLEWTDYKSKV 330
QY 366 FPOCVINETRLGNVYRFLHRKALKDVRKYGDIPSGMKVLEVISVHLDNSRYDQPNLF 425
DB 331 FTQHYVNETLRANITIGGIFRRATYDIDIKGTYTRKGMVVSFRVHNLNPEYKDAKTF 390
QY 426 NPMRMOQONGGASSGSGSFTWGNMMPGCGPRLCAGSELAKLEMAVFIHHLVLKFM 485
DB 391 NPMRMOQNSSEANP-----ANYTFPGGGRPCPGYELARVLSVFLRIYVRFEM 442
QY 486 ELAEDDQPPAFPPVDFPNCPLIRYSR 511
DB 443 VPAAEDKLVFPFTTRQKRYPIIVKR 468

```

```

RESULT 4
QY 09FX29 PRELIMINARY; PRT: 512 AA.
AC 09FX29;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE STENOID 22-ALPHA-HYDROXYLASE, PUTATIVE.
GN T9L24.44.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RA "Arabidopsis thaliana chromosome 1 BAC T9L24 genomic sequence.";
RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
RL EMBL: AC012396; AAG30983.1;
DR SEQUENCE 512 AA: 58297 MW: 1590C0304BF319FA CRC64;

```

Query Match 32.3%; Score 866; DB 10; Length 512;
 Best Local Similarity 33.8%; Pred. No. 1e-53;
 Matches 181; Conservative 105; Mismatches 173; Indels 76; Gaps 7;

```

QY 12 LLLPSSLLEFLTL-----LKRNRKTRFNLPPCKSGMPFLGELTIGLKYTATTLGD 66
DB 11 LLSVSSSTFLFAITIFLLAGIARRRRAPHRLLPSSRWPLGDFPALLNAVAGSHPS 70
QY 67 FMOQHSK-----YGRYSNLGFEPTIVSADAGLNRFIL 101
DB 71 FYEKQIKRKFVSLCSVLLILKRPDNGSREIRYGRIFSCSLFGKAAVYADPDKFRFTM 130
QY 102 ONEGRLEFCSYPRSIGGLGKWSMLVGDHMRMSISLNLSHARLRTLLKDOVERHT 161

```

```

DB 131 ONEGRLEFSSYPRKSFEDLVGKXGVITVHGDOQRRLHSIASSMMRHDLKTHFLVPIVYM 190
QY 162 LVFLDWSQOONSIFSADAEKAKFTFNIMAKHMSMDPEEETEOLKKEYTFMKGVASAPL 221
DB 191 LOTLSNFKDGEVALLDODIKRVAIHLMNQLGVS--SESEVDEMSQLFSDPDGCLSVPI 249
QY 222 NLPCTAHHKALQSAATILKFERKMEERKLDIKKEEQEVEKTEDEAEMSKSDHVRKOR 281
DB 250 DLPGFTYNAKAKRETIKINKTIEKRLONKAASD-----T 286
QY 282 TDDDLGWLKHSNSTEQILDLISLFRAGHTSSVALALAIFFLOACPRAVEELREH 341
DB 287 AGNGVLGRLEESLPNESMADEFTINLFRAGNETTSKTLFPAVYFLHCPRKMTQLEEH 346
QY 342 LEIARAKKEIGESLMWDYKKKDFLOCVINETRLGNVYRFLHRKALKDVRKYGDIPS 401
DB 347 -----DLRAGMGLTWQDKTMDPQCYIDETLRLGGIAIWLMPRAKDVSYODVPIK 399
QY 402 GKKVLPVIAVHLDNSRYDQPNLFNPMRW-----OOONGASSGSGSFTWGNMMPG 456
DB 400 GCFVYRFLSAVHLDSEYKESLSFNPWRKLDPEIQOKRMWRISP-----FYCPG 449
QY 457 GGPRLCAGSELAKLEMAVFIHHLVLKFMWELAEDDQPPAFPPVDFPNCPLIRYSR 511
DB 450 GGTRECPGAEIARLQIALFLHYFITTYKWTQLEKDRISFPSPARLVNGFKIQLNR 504

```

```

RESULT 5
ID 023242 PRELIMINARY; PRT: 524 AA.
AC 023242;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-AUG-1999 (TREMBlrel. 11, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CYTOCHROME P450 (CYTOCHROME P450 LIKE PROTEIN).
GN ROTUNDIFOLIA3 OR C7A10.980.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361880; PubMed=9694802;
RA Kim G., Tsukaya H., Uchimiya H.;
RT "The ROTUNDIFOLIA3 gene of Arabidopsis thaliana encodes a new member
RT of the cytochrome P-450 family that is required for the regulated
RT polar elongation of leaf cells.";
RT Genes Dev. 12:2381-2391(1998).
RN [2]
RP SEQUENCE OF 68-524 FROM N.A.
RA Bevan M., Terry N., Vos P., Heijnen L., Mewes H.W., Schueller C.,
RA Chaiwatzis N.;
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 270-524 FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Hohnsels J., Jesse T., Heijnen L., Vos P., Mewes H.W., Mayer K.,
RA Schueller C.;
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOOPROTEIN + O(2) = ROH +
CC OXIDIZED FLAVOOPROTEIN + H(2)O (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DB EMBL: AB008097; BAA37167.1;
DB EMBL: Z99708; CAB16850.1;
DB EMBL: AL022141; CAA18139.1;
DR InterPro: IPR001005;
DR InterPro: IPR001128;
DR Pfam: PF000067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.

```

DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
 KW Electron transport; Endoplasmic reticulum; Heme; Hypothetical protein;
 KM Membrane; Microsome; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 524 AA; 59389 MW; 550578908BDDF272 CRC64;

Query Match 32.2%; Score 864; DB 10; Length 524;
 Best Local Similarity 35.4%; Pred. No. 1.5e-53;
 Matches 186; Conservative 96; Mismatches 165; Indels 78; Gaps 10;

24 FLII-----LKRNRKTRFN-----LPPKSGMPFLEGTIGYL--- 56
 30 FLVTLGILLRPMILMFRLNRKSKTGDDEEDNEEKGMIPNGSLGWPEIGETLNFACG 89
 57 ---KPYATTLGDPMOQHVSKYKIRSNLGEPTIVSADAGLNRFLQNGRLLFECSYP 113
 90 YSSRPV-----FMDKRKSLYGKVFKNIIIGTPIIISTDAEVKAYVLQNGNFFVAPY 143
 114 RSIGGILGKSMVLVGDMDHMRKSLINLSHARLRTILIKDYERTTLVLDSSMOONSI 173
 144 KSTIELLENSILSINGPHOKRLHTLIGAFRLSPHLKDRITRDIASVLTSLASMAOLPL 203
 174 FSADAEAKKFFFNIMAKHIMSDPGEEETDQKKEVYFMKGVVSAPLNPGTAYHKALQ 233
 204 VHODEIKKMTFELIVKVMSTSPG-EDMNLTKLEFEETIGLICIPKPGTILYSLK 262
 234 SRATILKFERKMEERKLIDKEEDQEEVEYK-----EDEAKMSDHYAKKQRTDDL 287
 263 AKERILKVKVVEERQVAMTTTSPANDVYVLLRDGDSKQSPDSFVSGK----- 315
 288 GWVLKHSNLSTEOILDLISLIFAGHETSSVAIALAIFLOACRAVELREHLEIARA 347
 316 -----IVEMNIPGEETMTAMTLAVKFLSDNPALAKLVEENMEKRR 358
 348 KKEIIESELNMDYKKMDTCVINEETLRLGNVRFELRKALDKVRYGYDIPSGMKVLP 407
 359 KLELGE-EYKKTDMSLSTFQNVINETLRMANIINGWRKALKDVEIKGYILPQWCYLA 417
 408 VISAVHLDSNRYPDNLFNPRMOOQNGASSSGSSESTWGNVMPFGGRLCAGSEL 467
 418 SFISVHMEDEIYDNPQDFPWRMDRINGSANSSIC-----FTFPGGQRLCPGLEL 468
 468 AKLEMAVFHHLVLEKFMNELAEDDOPFAFPVDFPNCGLPIVSRIL 512
 469 SKLEISTFLHLVTRYSM-TAEDEIVSFPTVKKRRLLPIRVATV 512

RESULT 6
 09M066 PRELIMINARY; PRT; 457 AA.
 AC 09M066;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CYTOCHROME P450 LIKE PROTEIN.
 GN ATG36380.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RA PARTIAL SEQUENCE FROM N.A.
 RA Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
 RA Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN 121
 RP SEQUENCE FROM N.A.
 RA Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
 -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM (BY

CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL: AL161589; CAB80304.1; -
 DR InterPro: IPR001005; -
 DR InterPro: IPR001128; -
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PRO0385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
 DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
 KW Electron transport; Endoplasmic reticulum; Heme; Membrane; Microsome;
 KM Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 457 AA; 51607 MW; BD6A148E44EACD2 CRC64;

Query Match 32.1%; Score 860; DB 10; Length 457;
 Best Local Similarity 36.8%; Pred. No. 2.4e-53;
 Matches 179; Conservative 93; Mismatches 160; Indels 54; Gaps 8;

39 LPPKSGMPFLEGTIGYL-----KPYATTLGDFMOQHVSKYKIRSNLGEPTIVSA 92
 2 IPNGSLGMPVIGETLNFACGYSSRPVT-----FMDKRKSLYGKVFKNIIIGTPIIIST 55
 93 DAGNRFILONEGRLEFCYPRSTIGILGKSMVLVGDMDHMRKSLINLSHARLRTI 152
 56 DAEVKKVVLQNHGNTFVPAYRSTITELGENSILSINGPHOKRLHTLIGAFRLSPHLKDR 115
 153 LKLDVERHTLFLDSMOONSIFFSAODEAKKFFFNIMAKHIMSDPGEEETDQKKEVY 212
 116 ITRDIASVYVLTSLASMQPLVHODEIKKMTFELIVKVMSTSPG-EDMNLTKLEFE 174
 213 MKGVVSAPLNPGTAYHKALQASRTILKFTIRKMEERKLIDKEEDQEEVEYK-----E 266
 175 IKGLICIPKPGTILYSLKAKERLILKVKVVEERQVAMTTTSPANDVYVLLRDGSD 234
 267 DEAKMSKSDHYAKKQRTDDLLGWLKHSNLSTEOILDLISLIFAGHETSSVAIALAIF 326
 235 SEKQSPDSFVSGK-----IVEMNIPGEETMTAMTLAVK 270
 327 LQACPKAVEELREHLEIARAKKELGSELNMDYKKMDTCVINEETLRLGNVRFELR 386
 271 LSDNPVALAKLVEENMEKRRKLELGE-EYKKTDMSLSTFQNVINETLRMANIINGWR 329
 387 KALNDVRYKGYDIPSGMKVLPVISAVHLDSNRYPDNLFNPRMOOQNGASSSGSFS 446
 330 KALKDVEIKGYLIPKGCYVLAFLSVHDEDIYDNPQDFWRDRINGSANSSIC---- 385
 447 TWGNVMPFGGRLCAGSELALEMAVFIHHLVLEKFMNELAEDDOPFAFPVDFPNCGLP 506
 386 -----FTFPGGQRLCPGLELSKLEISTFLHLVTRYSM-TAEDEIVSFPTVKKRRLLP 439
 507 IRVSRIL 512
 440 IRVATV 445

RESULT 7
 09LNT3 PRELIMINARY; PRT; 478 AA.
 AC 09LNT3;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE T12C24.27.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RA PARTIAL SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shin P., Altafi H., Bel O., Chin C., Chio J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Hong B., Koo T., Lam B., Lee J.,

Db 330 GELGKLDIAFLHHLVLYSTRMKIKSDMPIAHVPERKGMLEI 374

RESULT 9

ID 09LIC5 PRELIMINARY; PRT; 464 AA.

AC 09LIC5; 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, last annotation update)

DE CYTOCHROME P450-LIKE PROTEIN.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OC NCBI_TaxID=3702;

ON 11

RP SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA:

RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA:

RX PubMed=10907853;

RA Nakamura Y.;

RT "Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the regions of 4,251,695 bp covered by ninety pl, TAC and BAC clones."

RT DNA Res. 7:217-221(2000).

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

DR EMBL: AP001307; BAB01922.1; -.

DR InterPro: IPR000504; -.

DR InterPro: IPR001128; -.

DR Pfam: PF00067; P450; 1.

DR PRINTS: PR00385; P450.

DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.

DR PROSITE: PS00030; RNP_1; UNKNOWN_1.

KM Heme; Monooxygenase; oxidoreductase.

SQ SEQUENCE 464 AA; 52864 MW; A65E09465E5190B CRC64;

Query Match 29.4%; Score 789; DB 10; Length 464;

Best Local Similarity 36.5%; Pred. No. 2.7e-48;

Matches 163; Conservative 101; Mismatches 145; Indels 38; Gaps 8;

QY 40 PPKSGMPFLEGTGYLKPYATTATLGDPMOAHVSKGIYRSNLFGEPTIVSADAGLNR 99

DB 52 PHGSLGMPVIGETIEFYSAVSADRPSFMDKRLMYGRVFSHIFGATIVSDAEVNA 111

QY 100 ILONEGRLEFESYPRISIGILGKMSLVLDGMDHMRKRSISLNLFLSHARLTLTKOVER 159

DB 112 VLQSDSTAFAFVPEYKTYRELMGKSSILLINGSLRRHFGVSLKSPLLAKQIVRMHK 171

QY 160 HTLVLDLSDMOONSIFSAODEAKKFTFNLMAKHMSMDPGEETDOLKEEYTFPKGVSA 219

DB 172 FLSSMOLMSDQVLLQDVSQYAFVYAKALISVEKG-EDLELKRPEENFISGLMSL 230

QY 220 PLNLPGTAHYKALOSRATILFKIERKMEERKLIDKEEDQEEVEYKTEDEAM--SKSD 276

DB 231 PINPGTQLHRSIQAKKNMVKQYERIEGKIR--KTKNKEEDVIADVDVLLKDSSEH 288

QY 277 VRKORTDDDLGWLKSNLSTEDILDLISLFLAGHETSSVAIALAIFLQACPKAVEE 336

DB 289 -----LTH-NLIANNMIDMMI---PGHDSVPVLLITLAVKFLSDSPALNL 329

QY 337 LREELHETAKRKEGSELNMDQYKKMDFTQCYINFTLRLGNVYRLHRAKALDVAYKG 396

DB 330 LTEENMML-KSLKELTGEPLVNDYLSLPFOKAVITETLRGNVITGVMRKAMADVIEIKG 388

QY 397 YDIPSGKVLPLVISAVALDHSRYDOPNLFNFMQOONNGSAGSGSFTWGNMNYMPFG 456

DB 389 YVDPKMGCFPLAVLSVHLDKIKYTESPYKFNMRMRQERDMNTSS-----FSPFG 436

QY 457 GGPRLCAGSELAKLEMAVFIIHHLVLPK 483

DB 437 GGRLLCPGLDLARLETYSVFLHVLVTRF 463

RESULT 10

ID 09FMA5 PRELIMINARY; PRT; 465 AA.

AC 09FMA5; 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, last annotation update)

DE CYTOCHROME P450.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OC NCBI_TaxID=3702;

ON 11

RP SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA:

RX MEDLINE=98290546; PubMed=9628582;

RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N., Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV. Sequence features of the regions of 1,456,315 bp covered by nineteen physically assigned pl and TAC clones."

RT DNA Res. 5:41-54(1998).

DR EMBL: AB009048; BAB08653.1; -.

SQ SEQUENCE 465 AA; 53767 MW; 0C00459C9C866D1F CRC64;

Query Match 27.1%; Score 726; DB 10; Length 465;

Best Local Similarity 32.3%; Pred. No. 8.4e-44;

Matches 166; Conservative 91; Mismatches 181; Indels 76; Gaps 10;

QY 12 LLLPLSLLLFLILIKRN--RKTFFNLPPKSGMPFLEGTGYLKPYATTATLGDPMQ 69

DB 11 LLIVSLCSALL-----RNNOMRYTKNGLPFGMGMPIGETIEFLKQSP-----NFM 59

QY 70 QHVSXGKIYRSNLFGEPTIVSADAGLNRFLIONEGRLEFESYPRISIGILGKMSMLV 129

DB 60 NQRLRYGSFFKSHLGCPTLISMDSEVNRITLKESGLVPGVQSLDLGTCNMAVH 119

QY 130 GDMHRKRSISLNLFLSHARLTLKDVERRHTLVLDSDMOONSIFSAODEAKKFTF---- 185

DB 120 GSHRLKRGSLSLISSSTMKRHDILPKVDHFMRSYLDQWMELEVYIDQTKHMAFLSL 179

QY 186 -----NLMAKHMSMDGEEETDOLKEEYTFPKGVSAFLNLPGTAYHKALOSRAITLK 240

DB 180 TQIAGNLRKPFV-----EEFKTAFPKLVAGTSLVPIIDLPGTIVYRCGIARNNIDR 229

QY 241 FIERKMEERKLIDKEEDQEEVEYKTEDEAKMSKSDHVRKORTDDDLGWLKSNL-----L 296

DB 230 LIRELMERR-----DSGE-----TFDMLGTYLAKKEGNNPFL 262

QY 297 STEQIIDLISLFLAGHETSSVAIALAIFLQACPKAVEELREELHETAKRKEGSEL 356

DB 263 TDEIRIQVYVITLYSGYETVSTSMALKYLHDPKALQELRAENLAFRRRRKQ--DEPL 320

QY 357 NMDDYKKMDFTQCYINFTLNLGNVYRLHRAKALDVAYKGYDIPSGKVLPLVISAVALD 416

DB 321 GLEBYKSMKFTTRAVIYETSLRATIVNGVLRKTRDELINQVLLIPKGRIRIVYTRREIYDA 380

QY 417 SRYDQPLFNFMQOONNGSAGSGSFTWGNMNYMPFGGPRLLCAGSELAKLEMAVFI 476

DB 381 NLIEDPLIFNFMKMSLESQ-----NSCVPFGGTRLCGKELGIVEISSFL 429

QY 477 HHVLKENWELAEDDQPAFPVDFPGLPIRVS 510

DB 430 HYFVTRRMEDIGDDELMAVFPVFAFPAKGFRLRIS 463

RESULT 11
ID 09LH81 PRELIMINARY: PRT: 465 AA.
AC 09LH81.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CYTOCHROME P450.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneo T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX PubMed=10907853;
RA Nakamura Y.;
RT *Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the regions of 4,251,695 bp covered by ninety P1, TAC and BAC clones.";
RL DNA Res. 7:217-221(2000).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AB002060; BAB02270.1; -
DR InterPro; IPR001128; -
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 465 AA: 53862 MW; 630A21D0765ED0D CRC64;

Query Match 26.5%; Score 711; DB 10; Length 465;
Best Local Similarity 32.1%; Pred. NO.9.8e-43;
Matches 163; Conservative 93; Mismatches 195; Indels 56; Gaps 8;
QY 14 LPLSLTLLFLTLK-----RRNRKTRFNLPGKSGMPFLGTTGYLKPYATTLGDMF 67
DB 3 IMMILGLVITVCTALLRNQMRYSKGLPPTGMGPPIGETTEFLAKGP-----DF 57
QY 68 MOQHSKYGKIRSNLFGPEPTVSADAGLNRFILQNEGRLEFECSPRSIGILGKMSLV 127
DB 58 MNQRLRYGSEFKSHILGCPITVSMQAEINRYILMESKGLVAGYQSMDLITGCNIAA 117
QY 128 LYGDMHQRMSISLNFSLHARLTLIKDYERHTLFVLDSMOONSIFSQODEAKKTFPL 187
DB 118 VGPSRLMRGSLSLISITPMKMDLPLKIDPMRNYLGGMDLETVDIOETKHMALF 176
QY 188 MAKHISMDPGEETEQLEKEVTFMKGVASAPLNPGTAHYKALOSRATILKFIKKME 247
DB 177 SSLQIAETLKRPVEBEYETFEFKLVGTLSPIDIPGINYSQGARNNIDRLLELMQ 236
QY 248 EKKLIDKEDEEVEEYKTEDEAEMSKSDHVRKORTDDLLGWLKHSN---LSTEQILD 303
DB 237 ERK-----ESGETF-----DMLGKKEEDNRRLITDKEIRD 269
QY 304 LLSLIFAGHETSSVAIALAIFFLQCPKAVEELREHELEIRAKKELGESELMNDYRK 363
DB 270 QVVTILYSYETVSTSMALXKTLHDHPALBELREHLEIRKRP--DEPLTLDIDS 327
QY 364 MDTQCVINETLRGNVRFRLHRAKALKDVRKGYDIPSGMKVLPVISAHLNDSRYDQN 423
DB 328 METRAVIEPETSRLATIVNGVLAKTJHDELNGYLIPKGMRIYVYREINYSIATEDDM 387
QY 424 LNPFRMOQONNGASSGSGSFSTGCMNNTMPFGGPRLCAGSELAKLEMAVFIHLVLP 483
DB 388 INPFRMMEKSLDESKS-----YFLFGGVRLCPGKRELISSEVSLHYEVTKY 436

QY 484 NMELAEDDQFAFPFVDPNGLPKIRVS 510
DB 437 RMEENGEDKLMVFPKVSAPKGYHKCS 463
RESULT 12
ID 09FH76 PRELIMINARY: PRT: 463 AA.
AC 09FH76.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CYTOCHROME P450.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneo T., Kato T., Asamizu E., Kotani H.,
RA Tabata S.;
RT *Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence features of the regions of 3,076,755 bp covered by sixty P1 and TAC clones.";
RL DNA Res. 7:31-63(2000).
DR EMBL; AB020744; BAB10255.1; -
SQ SEQUENCE 463 AA: 52366 MW; COD17293F53F812 CRC64;

Query Match 24.5%; Score 658; DB 10; Length 463;
Best Local Similarity 31.8%; Pred. NO.5.8e-39;
Matches 163; Conservative 95; Mismatches 184; Indels 70; Gaps 15;
QY 14 LPLSLTLLFLTLK-----RRNRKTRFNLPGKSGMPFLGTTGYLKPYATTLGDMF 68
DB 6 LPLTSSAALFLCLLFLTIGVRRSSSTKLPLPPTGWTGYVET--FLYSQDP-NVFF 61
QY 69 OQHSKYGKIRSNLFGPEPTVSADAGLNRFILQNEGRLEFECSPRSIGILGKMSLV 128
DB 62 AAKQRYGSEFKSHILGCPITVSMQAEINRYILMESKGLVAGYQSMDLITGCNIAA 121
QY 129 VGDHARDMSISLNFSLHARLTLIKDYERHTLFVLDSMOONSIFSAODEAKKTFPL 188
DB 122 QGDYSKRLKLVRAFMPDAIRN-MVPHIESIAOESLNSMDGTQLTNTYQ-EMKTYFNVA 179
QY 189 AKHISMDPGEET---EQLEKEVTFMKGVASAPLNPGTAHYKALOSRATILKFIKKME 245
DB 180 LISL---GKQEVYRREDLKRCYILLEKYNMPLNLTGTLFHRAMKARKELAQILANI 235
QY 246 MEERKIDKEEVEEYKTEDEAEMSKSDHVRKORTDDLLGWLK-HSNLSTEQILD 304
DB 236 LSKRR-----QNPSSNR-----DLGSMEDKAGLTLDQIADN 268
QY 305 LLSLIFAGHETSSVAIALAIFFLQCPKAVEELREHELEIRAKKELGESELMNDYRK 364
DB 269 IIGVIFAARDTASVLTWLKYLADNPVLEAVTEQMAIRDKKE-GES-LTWEDTKKM 326
QY 365 DFTQCVINETLRGNVRFRLHRAKALKDVRKGYDIPSGMKVLPVISAHLNDSRYDQN 424
DB 327 PLTVRIIQTLLAATLSTFEREADVEYEGTLIPGKMKVLPLEFNHNDISDPDEK 386
QY 425 FNPFRMOQONNGASSGSGSFSTGCMNNTMPFGGPRLCAGSELAKLEMAVFIHLVLP 484
DB 387 FDPSSRE-----VAPKPTFMPFGSGHSCGNELAKLEISVLLIHLTTKXR 433
QY 485 WELAEDDQ-----PFAFPVDPNGLPKIRVS 511
DB 434 WSIYVPSDGIQYCPALP---QNGPLIALER 461

```

RESULT 13
ID 065624 PRELIMINARY; PRT; 457 AA.
AC 065624;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CYTOCHROME P450.
GN T18B16.200 OR A74G19230.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Beves V., Rechmann S., Borkova D., Ansoorge W., Bancroft I.,
RA Mewes H.W., Mayer K., Scheller C.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Beves V., Rechmann S., Borkova D., Ansoorge W., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 131-457 FROM N.A.
RA Van Der Schueren J., Chuang Y.J., Aert R., Defoor E., Robben J.,
RA Vockert G., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: AL021687; CA116713.1; -.
DR EMBL: AL161550; CAB78925.1; -.
DR HSP; P33006; 1CPT.
DR InterPro: IPR001128; -.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ
SEQUENCE 457 AA; 52436 MW; 6105FB7C181E4F07 CRC64;

Query Match 23.6%; Score 632; DB 10; Length 457;
Best Local Similarity 31.8%; Pred. No. 4e-37;
Matches 151; Conservative 92; Mismatches 176; Indels 56; Gaps 12;

QY 13 LLLPSLLSLFLILKRRKTRPNLPKSGMPFLGETIGYLRKYATTLGDPMOOHV 72
DB 10 LFASLFLYFRCISORFGSSKILPPGTMGPVGET---FQLYSQDP-NVFFSQK 65
QY 73 SKYKIVRSNLFGEPTIYSADAGLNRFILONEGRLEFECYSRISGILGKMSMLVLGDM 132
DB 66 KRYSVKRTHTVLGCPCKWISSPEAKFVLTYSKSHLFKTFPASKERMLGKATFFHOCGY 125
QY 126 HAKIRKLVLRAPFPESTIN-WVPDIESIAODLSRW-EGTMINITYQEKTYTFVALLSI 183
DB 133 HRDKSISLNLFLSHARLTILKDVERTTLVLDSDWQNSIFSAODEAKKFTFLMAKHI 192
QY 193 MSMDGEEET---EOLKEVYTFMKGVYAPLNPGRVYHAKALOSRATILFKIERKEER 249
DB 184 F---GKDEVLYRDLKRCYILKEGYNSMPVNLPGTLFHSKARKKELSOILARIISER 239
QY 250 KLIDKEEDQEEVEKTEDEAKMSKSDHVRKORTDDLLG-VLKLHNSNSTEQIIDLITSL 308
DB 240 R-----QNGSSR-----NDLGSFMGKEELTDQIADNITIGV 272
QY 309 LFAHETSSVAIALAIFLQCPRAVELREHLEIARAKKELESELNMDYKNDPFTQ 368
DB 273 IFAARDTASVMSMLIKLALANPNVLAENVLEQNAI-RKDKESGS-LTWDTDKMKPLTS 330

```

```

QY 369 CVINETLRGLGNVFRLLRKALKDVRKGYDIPSGMKVLPVYSAVHLNDRYDNLNPMW 428
DB 331 RVIGETLRVASILSFTRAEAVEDVEYEGYLIPKCMKVLPRFNRIHNSADLFSNGKRDPS 390
QY 429 RWQOQNNGASSGSGSFSTGNNMPFGGGRCLCAGSELAKLEMAVYTHLVKLF 483
DB 391 RFE-----VAPKRNTEPMFGNGTHSCPGNELAKLEMSIMIHLLTTYK 432

RESULT 14
ID 09F138 PRELIMINARY; PRT; 518 AA.
AC 09F138;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CYTOCHROME P450-LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA.
RX MEDLINE=9397451; PubMed=10470850;
RA Kaneo T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
RT Sequence features of the regions of 1,011,550 bp covered by seventeen
RT P1 and TAC clones."
RL DNA Res. 6:183-195(1999).
DR EMBL: AB017064; BAB1064.1; -.
SQ
SEQUENCE 518 AA; 58877 MW; CCA3F74F4301563E CRC64;

Query Match 23.5%; Score 631; DB 10; Length 518;
Best Local Similarity 30.5%; Pred. No. 5.7e-37;
Matches 152; Conservative 101; Mismatches 190; Indels 56; Gaps 9;

QY 27 LKRRNKRTRPNLPKSGMPFLGETIGYLRKYATTLGDPMOOHVSKYKIVRSNLFGE 86
DB 64 LYRWSNPKCKNGKLRPGSMGLPIIGETCDFEPHOLYELSPVKKRMKLYGELPFTNIFGS 123
QY 87 PTIVSADAGLNRFILONEGRLEFECYSRISGILGKMSMLVLGDMHRDKSISLFLSH 146
DB 124 NTVALTEPDIIIEVFRQENKSEFVSYPDEAFVKGPKGENVFLKHGNIKHVQISLHIGS 183
QY 147 ARLEITILKDVERTTLVLDSDWQNSIFSAODEAKKFTF-NLAKKHMSMDPGEETPOL 205
DB 184 EALKKKMIGELDRVTEHLRSKANOGSFDKAEAVESYIMAHLPKIIISNLKPEYQAT-L 241
QY 206 KKEYVTFMKGVYAPLNPGRVYHAKALOSRATILFKIERKEER-----LDIKE 255
DB 242 VDNIMALGSEWSPPLTITLITSTYKFIARVAYLVQIVTKFRKKSREMGDFLDTMV 301
QY 256 EDQEEVEKTEDEAKMSKSDHVRKORTDDLLGVNLKHSNLSTEQIIDLITSLFACHET 315
DB 302 EEGEREDVIFNEES-----AINTLFAILVAKES 330
QY 316 SVAIALAIFLQCPRAVELREHLEIARAKKELESELNMDYK-KNDFQCYINET 374
DB 331 TTSVTSIAIKFLAENHAKALAEKREHAAIIQNRNGK-AGVSWEYRHQMTFTMVINET 389
QY 375 LRLGNVFRLLRKALKDVRKGYDIPSGMKVLPVYSAVHLNDRYDNLNPMWMOON 434
DB 390 LRANMAPIMYKRAVNDVEIKGITIPAGMIVAVIPRAVHRDALTAYENPLENPRMRWGE 449
QY 435 NGASSGSGSFSTGNNMPFGGGRCLCAGSELAKLEMAVYTHLVKFWELAEDDQPF 494
DB 450 ---LRSGSKTF-----MFGGCVROCVGAEFARLQISIFIHLLVTTYDPSLQDSESEFI 499

```


OY 495 APEVDFPNCGLPIRVSRII 513
 DB 500 RAPLFFPKGLPIKISQSL 518

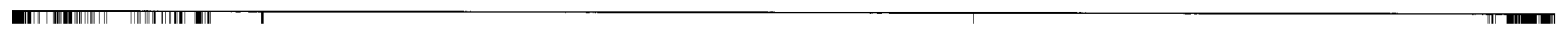
RESULT 15

09LG17 PRELIMINARY: PRT: 735 AA.
 AC 09LG17
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE F14J16.21.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
 Kim C., Altafi H., Bel Q., Chin C., Chou J., Choi E., Conn L.,
 Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B., Lee J.,
 Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
 Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 Ecker J.R.,
 RT "Genomic sequence for Arabidopsis thaliana BAC F14J16 from chromosome
 I.",
 RN Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.,
 RN Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.,
 RN Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RL [4]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.,
 RN Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RL [5]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.,
 RN Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RL [6]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 Khan S., Kim C., Altafi H., Bel Q., Chin C., Chou J., Choi E., Conn L.,
 Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B.,
 Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
 Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 Theologis A., Ecker J.,
 RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -I- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
 CC -I- OXIDIZED FLAVOPROTEIN + H(2)O (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM (BY
 CC SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL: AC002304; AAF79335.1; -.
 DR InterPro: IPR001128; -.
 DR Pfam: PF00067; p450; 2.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
 KW Electron transport; Endoplasmic reticulum; Heme; Membrane; Microsome;
 KW Monooxygenase; Oxidoreductase.
 SO SEQUENCE 735 AA; 84661 MW; 47186202974F811 CRC64;

Query Match 23.5%; Score 631; DB 10; Length 735;
 Best Local Similarity 29.8%; Pred. No. 9, 2e-37;

Matches 169; Conservative 99; Mismatches 187; Indels 112; Gaps 15;
 OY 19 ISLLFLILK-----RRNKRTPNLPFGKSGMPFLGETTGYLKPATTLTGPMQOH 71
 DB 190 LSFVIALVYVVKISLWLYRMANPNCSGKLPFGSGMGPVIGETVEFFKPSFNFHPEVKR 249
 OY 72 VSKY-----GKYSNLFGEPTIVSADAGNLRFTLQNGRFLFEGCSYRSIG 118
 DB 250 MKFLVPIYSAGKKNLHGSLFRTNLSKTTVSTPEVNFELTKQENRCFTIYSEALVR 309
 OY 119 ILGKSMVLVVG-DNHRDMRSISLNFSLHARLTILKDVERTFLVDSMOQNSIFSAQ 177
 DB 310 IFGKDNLFPGQKDFHRVWRHIALQLGPECLKQRFIQIDIAETSEHLKSVSFGGVQVK 369
 OY 178 DEAKKFTFNLAKKHMS-MDPEEETEOUKKEYVPMKGVASAP----- 220
 DB 370 DYSGRILIEOMTLMISNKP--ETRSKLIESFRDPSFPLVRSPPDPSPFNALYNGMLK 427
 OY 221 ----LNLPGTAYHKAOSRATILKPIERKMEERKLDIKEEDQEEBEVTEDEAKMSKD 275
 DB 428 DEVMILNIDPLFY--VKARSVMKMKRMKEREAREATSDSKTGDFMETIYEVK-- 483
 OY 276 HVKORTDDDLGWVLKHSNLSTEOILLISLPHAGHETSSVAIALAIFLQACPKAVE 335
 DB 484 -----EGDTINERSVELLSLLASVETSTMTALTVKFAIENPKVLM 527
 OY 336 ELREHELEIA--RAKKELGESLNDMDYK-MDFQCIVINELRLGNVYRFLHRAKADY 392
 DB 528 ELKREHETILQNRADKESG--VTWKEYRSMNPFHMVINESLRGSLSPAMFRKAVNDV 584
 OY 393 RYK-----GYDIPSGMKVLPVYSAVHLNDSRYDQPN 423
 DB 585 ELKGRFSLFCYIKITISLIPNDLIONRVVAGYITIPACMIYLVVPSLHYDPQIYEQC 644
 OY 424 LFNPMRWQOONNGASSSGSFTWGNNTYMPFGGPRCLCAGSELAKLENAVFTHLKYF 483
 DB 645 FENPMWECKE---LISGSKTF-----MARGGARLCAGAEFARLQMAIFLHILVTTY 694
 OY 484 NMELAEDDQPPAFPPVDFPNCGLPIRVS 510
 DB 695 DESLIDKSYITIRAPLRFK--PIRIT 719

Search completed: September 28, 2001, 18:45:40
 Job time: 130 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2001, 18:43:55 : Search time 13.15 Seconds

(without alignments)
1336.355 Million cell updates/sec

Title: US-09-502-426-2

Perfect score: 2681

Sequence: 1 MFETEHHTLPLLLPLSLLS.....FAFPVDFPNGLPFRVSRIL 513

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	978.5	36.5	472	1	C901_ARATH
2	761.5	28.4	464	1	CP85_LYCES
3	579.5	21.6	490	1	CP83_ARATH
4	503	18.8	492	1	CP26_BRARE
5	491.5	18.3	519	1	CP81_MAIZE
6	461	17.2	497	1	CP26_MOUSE
7	438	16.3	497	1	CP26_HUMAN
8	324.5	12.1	520	1	YRV2_CAEBL
9	321.5	12.0	518	1	YRV2_CAEBL
10	316	11.8	500	1	CP12_ARATH
11	312.5	11.7	504	1	CP32_RAT
12	312	11.6	501	1	CP15_MOUSE
13	310.5	11.6	520	1	YRV1_CAEBL
14	303	11.3	501	1	CP06_MOUSE
15	302	11.3	494	1	CPA8_MESAU
16	299.5	11.2	492	1	CPBC_RAT
17	297	11.1	491	1	CPB9_MOUSE
18	294.5	11.0	491	1	CPB4_RABIT
19	293	10.9	491	1	CPB1_RAT
20	291.5	10.9	530	1	CP51_YEAST
21	290.5	10.8	502	1	CP83_ARATH
22	289	10.8	503	1	CP39_RAT
23	288.5	10.8	492	1	CPB1_MOUSE
24	288.5	10.8	503	1	CP33_HUMAN
25	288.5	10.8	538	1	CP18_HUMAN
26	288	10.7	459	1	CP18_DROME
27	287.5	10.7	520	1	YRV5_CAEBL
28	287	10.7	506	1	CP92_ARATH
29	286.5	10.7	502	1	CP12_HUMAN
30	286.5	10.7	533	1	CP51_CANCA
31	284.5	10.6	505	1	C762_SOLME
32	281.5	10.5	504	1	CP3G_MOUSE
33	281	10.5	528	1	CP51_CANTR

34	280.5	10.5	504	1	CP3B_MOUSE	O64459 mus musculus
35	280	10.4	499	1	C771_SOLME	P37123 solanum mel
36	280	10.4	500	1	CP11_RABIT	P52786 oryctolagus
37	279.5	10.4	513	1	C773_SOYBN	O46928 glycine max
38	279.5	10.4	519	1	YRV8_CAEBL	O27520 caenorhabdi
39	279	10.4	501	1	C4D2_DROME	Q27589 drosophila
40	279	10.4	502	1	CPJ3_RAT	P51590 rattus norv
41	278.5	10.4	501	1	CP36_RABIT	P11707 oryctolagus
42	278	10.4	470	1	CPBK_MOUSE	O62397 mus musculus
43	278	10.4	503	1	CP3D_MOUSE	O64464 mus musculus
44	277.5	10.4	491	1	CPB5_RABIT	P12789 oryctolagus
45	276	10.3	491	1	CPB2_RAT	P04167 rattus norv

ALIGNMENTS

RESULT	ID	Query Match	Score	Length	ID	Description
1	C901_ARATH	36.5%	978.5	472	1	C901_ARATH
AC	042569:	Best Local Similarity	41.9%	Pred. No. 1.3e-55;		
DT	15-DEC-1998 (Rel. 37, Last sequence update)	Matches	210;	Mismatches	171;	Indels
DE	15-DEC-1998 (Rel. 37, Last annotation update)					
GN	CYP90A1 OR CYP90 OR CPD.					
OS	Arabidopsis thaliana (Mouse-ear cress).					
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;					
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;					
OC	Brassicales; Brassicaceae; Arabidopsis.					
OX	NCBI_Taxid=3702;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=CV. COLUMBIA;					
RX	MEDLINE=96200769; PubMed=8612270;					
RA	Szekeres M., Nemeth K., Koncz-Kalman Z., Mathur J., Kauschmann A.,					
RA	Altman T., Redei G.P., Nagy F., Schell J., Koncz C.;					
RT	"Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450,					
RT	controlling cell elongation and de-etiolation in Arabidopsis.";					
RL	Cell 85:171-182(1996).					
CC	-I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC	or send an email to license@sib-sib.ch).					
CC	-----					
DR	EMBL; X87367; CAA60793.1; -					
DR	EMBL; X87368; CAA60794.1; -					
DR	InterPro; IPR001128; -					
DR	Pfam; PF00067; P450; 2.					
DR	PROSITE; PS00086; CYTOCHROME_P450; 1.					
KW	Oxidoreductase; Monooxygenase; Heme.					
FT	BINDING 418 HEME (BY SIMILARITY).					
FT	SEQUENCE 472 AA; 53785 MW; 41A73F46D4E343F CRC64;					
QY	Query Match	36.5%	Score 978.5;	DB 1;	Length 472;	
QY	Best Local Similarity	41.9%	Pred. No. 1.3e-55;			
QY	Matches	210;	Mismatches	171;	Indels	41;
QY					Gaps	8;
QY	12 LLLPSLSLLFLILKRRNRKTRFPLPGSGMFLGEGTIGYIKPYATATLGDFMOOH 71					
QY	: : : : : : : : :					
DB	7 LLLSSIAAGFL--LLLRTRYRRMGRLPGSLGGLPIETQGLGAVYTEMPEPTIDR 63					
QY	72 VKRYGKIYNSNLFGEPTIYSADAGLNRFLLQNDGRFLFECYSRSGIGILGKSMVLVGD 131					
QY	: : : : : : : : :					
DB	64 VARYGSVFTHLFGEPITFSADPETRNRFVLQNEKLFECSPASYASICNLGKSHSLIMKS 123					
QY	132 MHRDMRISLNFSLSHARLRTILKQVERHTLFVLDSWQONSIFSAQDEAKKFTPNLMKMH 191					

```

      124 LAKRHSILYMSFANSSTIKHLMIDLVRFNLDSSSVRL--MEAKKTFELTVKQ 181
      192 IMSDPEEERHOLKKEVTFMKGVSAPLNPGTAHKAQSRATILKFERKMEERKL 251
      182 LMSDPC-EMSESRKEKELVIEGFSILPFLFSTYRKALQAR-----KVAEALT 232
      252 DIKEEDOEERKTEDEAEMSKSDHVRKQRTDIDLGVKLKHSNSTEQILDLISLFA 311
      233 VVVKRRREEEGAE-----RKQMLAALADGDFSDDEIYDFLALLVA 278
      312 GHESSVAIALAIFLOACPKAVELEREHLEIRAKKELGESLNDMDYKKMPTOCVY 371
      279 GYETSTIMLAVALFLEETPLALQKLEHEKIRAKMSD--SYSLWSDYKSMPTOCV 336
      372 NETRLGNVAFELHRAKLDVRYKGYDIPSGMKVLPVISAHLNDRYDQNLFPNRMQ 431
      337 NETLRVANIIGVRRAMTDEIKGYKIPKMKVFSFRAVHLDPNHFKAARTNPNRMQ 396
      432 QONNGASSSGSFSFTWGN--YMPGGGPRLCAGSELAKLEMAVFHHLVLKFMELAE 490
      397 -----NSVYTPSPNVFTPFGGPRLCAGSELARVALSVFLHRLVYGFMSVPAEO 446
      491 DQPAFPVDFPNDLPYRVS 511
      447 DKLVFPFTTQKRYPIFKR 467

```

RESULT 2

CP85_LYCES STANDARD; PRT; 464 AA.

```

AC Q43147;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE CYTOCHROME P450 85 (EC 1.14.-.-) (DWARF PROTEIN).
GN CYP85 OR D.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-CV. GCR758;
RX MEDLINE=96266705; Pubmed=8672892;
RA Bishop G.J., Harrison K., Jones J.J.G.D.;
RA "The tomato Dwarf gene isolated by heterologous transposon tagging
RT encodes the first member of a new cytochrome P450 family."
RL Plant Cell 8:959-969(1996).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U54770; AB017070.1; -
DR InterPro; IPR001128; -
DR Pfam; PF00067; P450.1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KM Oxidoreductase; Monooxygenase; Membrane; Heme.
FT BINDING 414 414 HEME (BY SIMILARITY).
SQ SEQUENCE 464 AA; 53706 MW; D2B21AAB7B14E94 CRC64;

```

Query Match 28.4%; Score 761.5; DB 1; Length 464;
 Best Local Similarity 33.4%; Pred. No. 9.5e-42;
 Matches 170; Conservative 90; Mismatches 188; Indels 61; Gaps 10;

```

      12 LLLPSLLSLFLFLILKRRN--RRTRENLPRGCKSPFLGTYGLPYATYATG-DFM 68
      5 LIPLSFFGLCIFTALLRMNQVAYKNLPRTGIMGWLFBETTEFLK-----LGSPFM 58
      69 QOHVSKTKIYRSNLFGEPPTVSADAGINFTLONEGRLECSYPRSIGILGWSMLVL 128
      59 KNQARVGSFEKSHILCCPTIVSMDSELNRYILVNEAKGLVPGYOSIDILGKCNIAAY 118
      129 VGDHRRMRSISLNFSLHARLRTLLKDVREHNTLFVDSMQONSIFSADAEKFTFLM 188
      119 NSGAKTKMRGALLSLISPTMIRDLPRKIDFMSHLTN--DNKVIDIQETNKAFLSS 177
      189 AKHIMSNDPEERT--EOLKREYVTFMKGVSAPLNPGTAHKAQSRATILKFERK 245
      178 LKQI-----AGIESISLQDFSEFFNLVLTSLPINLPNTNHYHGRQARKIYNLLRTL 233
      246 MEERKLDKEEDOEERKTEDEAEMSKSDHVRKQRTDIDLGVKLKHS-----NLSTEQI 301
      234 IEERR-----ASKETQHDMLGYLMEAEATRFKLTQDEM 266
      302 LDLLSLFLAGHETSVAIALAIFLOACPKAVELEREHLEIRAKKELGESLNDMDY 361
      267 IDLIITILYSGYETVSTSMVAAYKLDHDPVLELKEHNAIREKKR--EDPIDNDY 324
      325 RSMFTRAVILLETSLRATIYNGVLKRTQDMEINGIILPKGMRIYVYTRELNDYPRLYPD 384
      362 KKMDFTQCVINETRLGNVAFELHRAKLDVRYKGYDIPSGMKVLPVISAHLNDRYDQ 421
      422 PNLFPNRMQOONNGASSSGSFSFTWGN--YMPGGGPRLCAGSELAKLEMAVFHHLVL 481
      385 PYSFPNRMMDKS-----LEHONSFLVGGGTROCPGKELVAGVETSLHAFVT 433
      482 KFMNELAEDDQPAFPVDFPNDLPYRVS 510
      434 KYRWEELGGDKLMKFPYEAENGRLIRVS 462

```

RESULT 3

C883_ARATH STANDARD; PRT; 490 AA.

```

AC 023051;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE CYTOCHROME P450 88A3 (EC 1.14.-.-).
GN CYP88A3 OR YUP8H12.23.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-CV. COLUMBIA.
RA Theologis A., Osborne B.I., Vysotskaya V.S., Federspiel N.A.,
RA Toriumi M., Yu G., Oji O., Araujo R., Chung E., Dewar K., Dietrich F.,
RA Ecker J.R., Marziani A., Oelner P., Davis R.W.;
RA Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AC000098; AB071462.1; -
DR InterPro; IPR001128; -
DR Pfam; PF00067; P450.1.
DR PRINTS; PR00385; P450.

```


QY 289 WLK-HSNLSTEOILDLISLAFAGHETSVAIALAIFLOACPKAVEELREHLEIARA 347
 DB 303 EVLAKVDRKLTDEITIGOLFVFLAGYDTALTALSSSYLLATLHPICIKLOEE-----V 356
 QY 348 KKLGESELMNDYKMKDEFQCVINETLRGNVVRFLH-RALKADVRKYKDYDIPSGMKVL 406
 DB 357 DRECPDEPVEFTDOLSKLTYECLVCEKALRYPLASLVHNRKCLKTPNVLGMEIEAGTMIN 416
 QY 407 PVISAVHLDNRSY-DQPLNLFPMRQOONNAGSSGSGSFSTGNNVMPFGGPRCLAGS 465
 DB 417 VDTSLHNDPRVWGDVVEFKPERME-----SGDELFFPAKG-GYLPFGKGPRICTGM 467
 QY 466 ELAKLEMAVFIHLHLKFNWE 486
 DB 468 RLAMEMKMLJTNILKNYTFE 488

RESULT 10
 C912_ARATH STANDARD; PRT: 500 AA.
 AC 065790: O9S2U3;
 DF 01-OCT-2000 (Rel. 40, Created)
 DF 01-OCT-2000 (Rel. 40, Last sequence update)
 DF 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CYTOCHROME P450 91A2 (EC 1.14.-.-).
 GN CYP91A2 OR AT4G37430 OR F6G17.80.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_COLUMBIA; TISSUE=Seedling;
 RX MEDLINE=96281573; PubMed=9620263;
 RA Mitsuami M., Ward E., Ohta D.;
 RT "Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of
 RT cDNAs, differential expression, and RFLP mapping of multiple
 RT cytochromes P450.";
 RL Plant Mol. Biol. 37:39-52(1998).
 RN 12
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_COLUMBIA;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Mambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entlan K.-D., Terryn N.,
 RA Harris B., Ansoyge W., Brandt P., Grivell L., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohenseil J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van den Schueren J., Glymonprez B., Chuang Y.-C., Vandebussche F.,
 RA Breken M., Welfens I., Voet M., Bastiaens I., Aert R., Deteor E.,
 RA Wetzsteiger T., Boche G., Ransperger U., Hilbert H., Bruun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dikse W.,
 RA Moeliman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Benelaiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysnaert C., Giejen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
 RA Pettett A., Rajandream M.-A., Lyne M., Benes V., Reichmann S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehnert T.-H.,
 RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fattmann B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Argitidou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenet O., Outley F., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedid F., Cooke R., Berger C., Monfort A., Casauberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,

RA Zaccaria P., Beyan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Pannell L., Dedhia N., Gnoj L., Schütz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Hamon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Mix P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Krimer J., Fulton L., Mardis E., Dante M., Pepin C., Hillier L.,
 RA Nelson J., Spleth J., Ryan E., Andrews S., Giesel C., Layman D.,
 RA Du H., Ali J., Bergoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekhar M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
 RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:769-777(1999).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D78607; BAA28539.1; -;
 DR EMBL: AL035601; CAB38210.1; -;
 DR EMBL: AL161591; CAB80408.1; -;
 DR Mendel: 29894; Arath:1113.29894.
 DR InterPro: IPR001128; -;
 DR Pfam: PF00067; P450.1;
 DR PROSITE: PS00086; CYTOCHROME_P450.1;
 KW Oxidoreductase; Monooxygenase; Membrane; Heme.
 FT BINDING 438
 FT CONFICT 106 A->V (IN REF. 1).
 FT CONFLICT 127 I->M (IN REF. 1).
 FT CONFLICT 140 N->I (IN REF. 1).
 FT CONFLICT 454 454. S->T (IN REF. 1).
 FT SEQUENCE 500 AA; 57555 MW; 0FB453D2070EA2EA CRC64;
 SQ

Query Match 11.8%; Score 316; DB 1; Length 500;
 Best Local Similarity 25.2%; Pred. No. 3.1e-13;
 Matches 135; Conservative 83; Mismatches 199; Indels 118; Gaps 22;

QY 9 LPLLLPSLSLFLILKLR-RNRKTRFNLPGKSGMPLGETIGYKPYATATLGD 66
 DB 1 MLYFILP-----LFLVSYKFLSKYQRFNLPGPSPRPVGH-LHLMKP-----PIHR 50
 QY 67 FMOQHSYKGYKTRNSLNGEPYIV-----SADAGLNRIILONEGRFLFCSPRSIG 117
 DB 51 LLQRYSNQYGLIF-SLRFSSRRVYVITSPSLAOESTGONDVLSSRPLQTLAKVAYVNH 109
 QY 118 GLTGKMSMLVGLGDHNRDRI-SLNFSLHARLRT--LTKDVERHTFLVDSMOONSIF 174
 DB 110 TTVG-----TAPYGDWRNRLRICSOEILSSHLINPQHRKDEILMLTRLRLRYQTS-- 163
 QY 175 SAQDEAKRTF-----NLAKHIMSMPG-----EETEOIKREYVTFMKGV 216
 DB 164 ---NSNDPHTHELPLSLDPLFNNIVRWVTKRYRGDGVNKKKEALFKK----- 211
 QY 217 VSAPLNLPGTAVHKLQSRAT-----LKTIERKMERKRLDIEEDOEDEEVKTEDEA 269
 DB 212 -----LVYDIAMYSQANHSADYLPRLKLFKGKFEVYVAIG-----KSMDDI 253
 QY 270 ESKSGDHVKKORTDLDGLGVKHSNLSLSTEOILD-----LISLAFAGHETSVAIALAI 324
 DB 254 LKRLDLECRKRGKGNVNHHLISLDOOQPEYTTDVIKILMSMLAGETSAVLENNAM 313
 QY 325 FFLQACPKAVEELREHLEIARAKKELGSEL-NMDDYKMKDFTQCVINETLRGNVVR 383

Db 314 ANLRNEVELEKARSE-----IDKIGKORLIDESIDIANVPIQNVSEFRLFPVAPF 367
QY 384 L-HRKALKDYRYKGYDIPSGMKVLPVISAHLDNSRDYDNLFPNPMQOONNGASSGS 442
Db 368 LIPRSPDMDKIGGYDVPRIYVWNAIAHROPEIMEPEEKFPDXY---NDGC----- 419
QY 443 GSFSTWNNY-----MFGGGRCLCAGSELAKLEMAVFIHHLVLFKFWELAEDDQ 492
Db 420 -----GSDYVYVKLMPFGNGRRTCPAGIGORIVTLATLSLIOCFFEMENKGE 468
RESULT 11
CP32_RAT STANDARD: PRT: 504 AA.
ID CP32_RAT P05183; Q64672; Q64629;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE CYTOCHROME P450 3A2 (PC 1.14.14.1) (CYP11A2) (P450-PCN2) (P450/6-
BETA-A) (TESTOSTERONE 6-BETA-HYDROXYLASE).
GN CYP3A2 OR CYP3A-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87064606; PubMed=3785219;
RA Gonzalez F.J., Song B.-J., Hardwick J.P.;
RT "Pregnenolone 16 alpha-carbonitrile-inducible P-450 gene family: gene
conversion and differential regulation";
RL Mol. Cell. Biol. 6:2969-2976(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RC MEDLINE=91254339; PubMed=2043144;
RA Miyata M., Nagata K., Yamazoe Y., Kato R.;
RT "A gene structure of testosterone 6 beta-hydroxylase (P45011A).";
RL Biochem. Biophys. Res. Commun. 177:68-73(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RC MEDLINE=95070139; PubMed=7979376;
RA Miyata M., Nagata K., Shimada M., Yamazoe Y., Kato R.;
RT "Structure of a gene and cDNA of a major constitutive form of
testosterone 6 beta-hydroxylase (P450/6 beta A) encoding CYP3A2:
comparison of the cDNA with P450PCN2";
RL Arch. Biochem. Biophys. 314:351-359(1994).
CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
ACIDS, AND XENOBIOTICS.
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
OXIDIZED FLAVOPROTEIN + H(2)O.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, ENDOPLASMIC RETICULUM.
CC -1- INDUCTION: BY PREGNENOLONE 16-ALPHA-CARBONITRILE (PCN).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M13646; AAAA1051.1; -
CC EMBL: X79319; CAAS5887.1; -
CC EMBL: X79320; CAAS5888.1; -
CC EMBL: U09742; AAAB2168.1; -
CC EMBL: U09734; AAB60492.1; -
CC EMBL: U09725; AAB60492.1; JOINED.

DR EMBL: U09726; AAB60492.1; JOINED.
DR EMBL: U09727; AAB60492.1; JOINED.
DR EMBL: U09728; AAB60492.1; JOINED.
DR EMBL: U09729; AAB60492.1; JOINED.
DR EMBL: U09730; AAB60492.1; JOINED.
DR EMBL: U09731; AAB60492.1; JOINED.
DR EMBL: U09732; AAB60492.1; JOINED.
DR EMBL: U09733; AAB60492.1; JOINED.
DR PIR: A25222; A25222.
DR HSSP: P14779; 1FAG.
DR InterPro: IPR001128; -
DR InterPro: IPR002397; -
DR InterPro: IPR002401; -
DR InterPro: IPR002402; -
DR InterPro: IPR002403; -
DR Pfam: PF00067; P450.1.
DR PRINTS: PR00385; BP450.
DR PRINTS: PR00463; BP450I.
DR PRINTS: PR00464; BP450II.
DR PRINTS: PR00465; BP450IV.
DR POSITE: PS00086; CYTOCHROME P450; 1.
DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT BINDING 443 443 HEME.
FT CONFLICT 56 57 GL -> AV (IN CAAS5887).
FT CONFLICT 429 429 D -> H (IN REF. 1).
FT CONFLICT 445 445 G -> D (IN REF. 1).
SQ SEQUENCE 504 AA; 57731 MW; C2C84AF736035AB2 CRC64;

Query Match 11.7%; Score 312.5; DB 1; Length 504;
Best local similarity 25.5%; Pred. No. 5, 2e-13;
Matches 137; Conservative 96; Mismatches 211; Indels 93; Gaps 24;

QY 9 LRLPULLPSSLISLLPILILK--RNRKTRFNLP--PGKGMFPGICTGYLRYTFITL 64
Db 3 LLSALTLTEWLVALLVILLYLRLGTHRHGIFKKOGIPRPKPLFGLVLYNR----GL 57
QY 65 GDFMQHVSRYKIRYSNLSNLEPTIVSA--DAGLNRFILQNEGRLEPC---SYPSISG 117
Db 58 GRDMDEKRYKYGKIW--GLFDGQTPVPAIMDTMIKIVL-----VRCFSVFVNRNRFGR 110
QY 118 -GLLKSMVLVLDGMHRMRSISLNFSLHARLRTILKQVNRHTLVLDLSMOONS---- 172
Db 111 VGIMKAVSVAKDEEMKRYRALLSPFTS--GRUKE--MFLIEQYGDILVYVYLQDAETGR 168
QY 173 -----IFSA--ODEAKKFTFNMAKHIMS--MDPGEETQOLK--EYVFMKGVYAPL 221
Db 169 PVIMKAVYFGVSMIDVITTSFGVNVDSLNPKDPFEYKIKLRFDFDFLFSVLYFP 228
QY 222 NLRGTAVHKA-----QSRATILKFIERKMEERKLIDNEEDQEEVTEDEAEMSKS 274
Db 229 LTP---IYELNLCIMPRKDSIAFFQKVRH--IKETRLDSNKHNR-----VDFQLMLN 277
QY 275 DHRKQRTDDDLGLWLNKSLSTEOILDLILSLFRAGHTSSVAALALAFIQAQKAV 334
Db 278 AH---NNSKDE-----VSHNALSVDLIAOSVIFRAGYVTTSTTSFVLYFLATHTDIO 329
QY 335 EEIRREHLEIARAKKELGESELMMDYKKMDFTQCVINFTLRIGNVYRFLHRAKLDVRY 394
Db 330 KRIQEE-IDCALRSK-----APPTDYIMEHEYLDWNLNETRLRPLIGNRLEKCKADIEL 384
QY 395 KGYDIPSGMKVLPVISAHLDNSRDYDNLFPNPMQOONNGASSGSFSFGNNYMP 454
Db 385 DGLFIRKGSVVTIPRYALHHDHPHMKPRPEFHRERSKEKGSIDP-----YYLTP 455
QY 455 FGCGPRCLCAGSELAKLEMAVFIHHLVLFKFWELAEDDQPAFFVDPNGLPIRVSR 511
Db 436 FGNGPRCIGMRPALNMKILATKVLQNSFQPKETQ-----IPLKLSR 480

RESULT 12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2001, 18:43:30 : Search time 19.47 Seconds
(without alignments)
2007.064 Million cell updates/sec

Title: US-09-502-426-2

Sequence: 1 MFETEHHTLLPLLPISLLS.....FAFPVPDPNGLPVRSRL 513

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2677	99.9	513	2 T46143	steroid 22-alpha-h
2	978.5	36.5	472	1 S55379	cytochrome P450 cy
3	866	32.3	512	2 H96759	probable steroid 2
4	860	32.1	457	2 D85429	cytochrome P450 1i
5	815	30.4	382	2 T48613	hypothetical prote
6	761.5	28.4	464	2 T07859	cytochrome P450 ho
7	632	23.6	457	2 T04444	cytochrome P450 -
8	611	22.8	482	2 T02739	probable cytochrom
9	581	21.7	487	1 C71417	cytochrome P450 d1
10	579.5	21.6	485	2 A84859	probable cytochrom
11	579.5	21.5	490	2 H8185	hypothetical prote
12	577	20.4	455	2 T48973	cytochrome P450-1i
13	547.5	18.6	489	2 D96813	probable cytochrom
14	497.5	18.3	460	2 D96813	hypothetical prote
15	491.5	18.3	519	1 T02263	cytochrome P450 DW
16	489.5	17.9	444	1 S75761	cytochrome P450 ho
17	479.5	13.8	349	2 A86329	hypothetical prote
18	369	13.3	518	2 T20908	hypothetical prote
19	357	12.6	453	2 C83722	cytochrome P450 hy
20	338	12.3	517	2 T20907	hypothetical prote
21	329.5	12.1	520	2 T24778	hypothetical prote
22	324.5	12.0	518	2 T24783	hypothetical prote
23	321.5	11.8	500	2 T04737	cytochrome P450 ho
24	316	11.6	520	2 T24777	hypothetical prote
25	310.5	11.6	500	2 T52175	cytochrome P450 mo
26	310	11.4	504	2 A25222	cytochrome P450 3A
27	306.5	11.3	491	2 S31277	cytochrome P450 2B
28	302.5	11.2	491	2 S31277	testosterone 16a-h
29	300	11.2	491	2 S31277	testosterone 16a-h

30	299.5	11.2	492	2 S27160	cytochrome P450 2B
31	297	11.1	491	2 A31047	testosterone 16alp
32	297	11.1	494	2 A33293	cytochrome P450 2A
33	294.5	11.0	491	1 O4RBP	cytochrome P450 2B
34	294	11.0	506	2 D96672	probable Cytochrom
35	294	11.0	516	2 T48140	flavonoid 3',5'-hy
36	292	10.9	491	1 O4RBP	cytochrome P450 2B
37	291.5	10.9	491	1 S35666	cytochrome P450 2B
38	291.5	10.9	530	1 A27491	lanosterol 14alpha
39	290.5	10.8	502	2 T05246	cytochrome P450 mo
40	288.5	10.8	504	2 A29410	cytochrome P450, g
41	287.5	10.7	520	2 T24780	hypothetical prote
42	287	10.7	503	2 J04702	cytochrome P450 3A
43	284.5	10.6	501	1 S38534	cytochrome P450 76
44	284	10.6	501	2 T04735	cytochrome P450 ho
45	282	10.5	491	2 I49625	testosterone 16a-h

ALIGNMENTS

RESULT	1	
T46143	steroid 22-alpha-hydroxylase (DMF4) - Arabidopsis thaliana	
N:Alternate names: protein T3A5.40		
C:Species: Arabidopsis thaliana (mouse-ear cress)		
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 16-Feb-2001		
C/Accession: T46143		
R/Biocheck: H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quettler, F.; S		
A:Submitted to the Protein Sequence Database, December 1999		
A:Reference number: 223024		
A:Accession: T46143		
A>Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-513 <BL0>		
A:Cross-references: EMBL:AL132979		
A:Experimental source: cultivar Columbia; BAC clone T3A5		
C:Genetics:		
A:Map position: 3		
A:Introns: 74/2; 182/3; 233/3; 338/3; 369/3; 396/1; 432/3		
A>Note: T3A5.40		
C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology		
C:Keywords: heme; iron; metalloprotein		
F:308-484/domain: cytochrome P450 homology <P45>		
F:462/Binding site: heme iron (Cys) (axial ligand) #status predicted		
Query Match	99.9%; Score 2677; DB 2; Length 513;	
Best Local Similarity	99.8%; Pred. No. 9.2e-166;	
Matches	512; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MFETEHHTLLPLLPISLLSLLFLILKRRNRKTRFLPPGKSGWPLGTTGYLKYT 60	
DB	1 MFETEHHTLLPLLPISLLSLLFLILKRRNRKTRFLPPGKSGWPLGTTGYLKYT 60	
QY	61 ATTTGDFMQHVSXGKTKYRSNLFGEPTVSADGLNFTLONGRRLPESYPSIGIL 120	
DB	61 ATTTGDFMQHVSXGKTKYRSNLFGEPTVSADGLNFTLONGRRLPESYPSIGIL 120	
QY	121 GKWSMLVGVGMHMDRMSISINFLSHARLRTILKDYERHTLFYVDSWQMSIFSADDEA 180	
DB	121 GKWSMLVGVGMHMDRMSISINFLSHARLRTILKDYERHTLFYVDSWQMSIFSADDEA 180	
QY	181 KKFENLMAKIMSMDPEEETDOLKKEYVTFPMKGVSAPIPLDGTAYHKAQSRAITLK 240	
DB	181 KKFENLMAKIMSMDPEEETDOLKKEYVTFPMKGVSAPIPLDGTAYHKAQSRAITLK 240	
QY	241 FIERKMERKLDIKEDDEEVEVTEDEAKMSGNHVKKOTDDLLGWLKHSNLSTEQ 300	
DB	241 FIERKMERKLDIKEDDEEVEVTEDEAKMSGNHVKKOTDDLLGWLKHSNLSTEQ 300	
QY	301 ILDLISLLFGHETSVAIALAIFLQACPKAVEELREELLETARAKKEIGESLNMDD 360	
DB	301 ILDLISLLFGHETSVAIALAIFLQACPKAVEELREELLETARAKKEIGESLNMDD 360	

Db 301 ILDLISLFLFAGHETSSVAIALAIFFLQACPKAVEELREELHETARAKKEGSELNMD 360
QY 361 YKMDFTQCYINFTLRNGVNFRLHRRKALDKVYKGYDIPSGMKVLPVISAHLNDSRYD 420
Db 361 YKMDFTQCYINFTLRNGVNFRLHRRKALDKVYKGYDIPSGMKVLPVISAHLNDSRYD 420
QY 421 QPNLFNPMWQOONNGASSSGSFSTWGNMYPFGGPRLCAGSELAKLEMAVFIHLV 480
Db 421 QPNLFNPMWQOONNGASSSGSFSTWGNMYPFGGPRLCAGSELAKLEMAVFIHLV 480
QY 481 LKFNWELAEEDOPFAFPVDFPGLPIRVSRL 513
Db 481 LKFNWELAEEDOPFAFPVDFPGLPIRVSRL 513

RESULT 2

S55379

cytochrome P450 CYP90 - Arabidopsis thaliana

N:Contains: oxidoreductase (EC 1.-.-.-)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C:Accession: S55379

R:Sequences: M.; Nemeth, K.; Koncz, Z.; Nagy, F.; Koncz, C.

Submitted to the EMBL Data Library, May 1995

A:Reference number: S55379

A:Accession: S55379

A:Molecule type: mRNA

A:Residues: 1-472 <S>E>

A:Cross-references: EMBL:X87367; NID:g853718; PIDN:CAA60793.1; PID:g853719

C:Genetics:

A:Gene: CYP90

C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology

C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;

F:775-440/Domain: cytochrome P450 homology <P450>

F:418/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 36.5%; Score 978.5; DB 1; Length 472;

Best Local Similarity 41.9%; Pred. No. 5,9e-56;

Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;

QY 12 LLLPSSLFLFLILKRRNRKTRFNLPPKSGWPFLEGTIGLYKPYATTLGDFMOOH 71
Db 7 LLLSLSTIAAGFL--LLRRTYRRKMGGLPPSGGLPLIGTFOLIGYKTENPEPFI 63
QY 72 VSKYKGYRNLGEPPIVSADAGLNKFLIQNGRLEFECSPRSIGIGLKWSMLVVG 131
Db 64 VARYGVFMTHLGEPTIFESADPETNRFVLQNGKLFECSPASICNLGKHSLLMKGS 123
QY 132 MHRDMSISINFLSHARLTILKDVBRHTLFLVDSMOONSIFSADDEAKKFPNLM 191
Db 124 LHRHMSLTMFSANSLIKHMLMDLRLFMIDSSSVLL--MEAKRTIFELTVKQ 181
QY 192 IMSNDPEEETDOLKREYVFMKGVVASPLNPGTAVHKLQSRATILKIERKMEER 251
Db 182 LMFDFRG-EMSESLRKYELVIGFSLPLPLSTYTRKAIQAR-----KKVAELT 232
QY 252 DIKEEDQEEBEVKTDEDAEMSKSDHVRKQRTDGLGWLKHSNLSTEOILLILSL 311
Db 233 VVVMKRRREEEGAE-----RKMDLAALLAADGFSDEEIVDFVALVLA 278
QY 312 GHETSSVAIALAIFFLQACPKAVEELREELHETARAKKEGSELNMDOKKMDFTQ 371
Db 279 GYETSTIMTLAKFLLETFLAQLKEEHEKIRAKSD--SYSLWSDYKSMPEQCV 336
QY 372 NETLRLGNVNFRLHRRKALDKVYKGYDIPSGMKVLPVISAHLNDSRYDQPNLF 431
Db 337 NETLRAVNIIGVFRAMTVEIKGYIKPKGVVSSFRVHLDPNHFKARFENPMRW 396
QY 432 QOONNGASSSGSFSTWGNM--YMPFGGPRLCAGSELAKLEMAVFIHLVLAKE 490
Db 397 -----SNSVTTGSPSNVFTPPFGGPRLCAGSELAKLEMAVFIHLVLAKE 446

QY 491 DOPFAFPVDFPGLPIRVSRL 511
Db 447 DKLVFEPTRTRQKRYPLFPKR 467

RESULT 3

H96759

probable steroid 22-alpha-hydroxylase T9L24.44 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: H96759

R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: H96759

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-512 <STO>

A:Cross-references: GB:AE005173; NID:g11120803; PIDN:AG30983.1; GSPDB:GN00141

C:Genetics:

A:Gene: T9L24.44

A:Map position: 1

Query Match 32.3%; Score 866; DB 2; Length 512;

Best Local Similarity 33.8%; Pred. No. 1.2e-48;

Matches 181; Conservative 105; Mismatches 173; Indels 76; Gaps 7;

QY 12 LLLPSSLFLFLILKRRNRKTRFNLPPKSGWPFLEGTIGLYKPYATTLGDFMOOH 66
Db 11 LLSVSSSTFLFLITLFLAGIARRRRRAHPRPLPSRGWPLGDFWAMNVAAGSPSS 70
QY 67 FMOOHVSK-----YGRYRNLGEPPIVSADAGLNKFL 101
Db 71 FVKKQIKKTFYSLCLSVLLILKRPDNGFNEIRYKGLFSCSLGKNAVVSADDFRFTM 130
QY 102 QNEGRLEFECSPRSIGIGLKWSMLVVGDMHRDMSISINFLSHARLTILKDVBRHT 161
Db 131 QNEGKLFQSSYPRKSPFDLVGCDGVTYHGDQRLHSIASMMRHQOLKTHFEVLPV 190
QY 162 LFLVDSMOONSIFSADDEAKKFPNLMMAKHMMDGGEETDOLKREYVFMKGVVASPL 221
Db 191 LQTLSPNKDGEVYLLQDICKKVAIHLMVNOLGV--SESEVDEMQLFSDPVGCLSP 249
QY 222 NLPGTAVHKLQSRATILKIERKMEERKLDIKEEDQEEBEVKTDEDAEMSKSDHVR 281
Db 250 DLPGFTYNNAMKARKKEIRINKTIERKLNKAAKD-----T 286
QY 282 TDDDLGNVLKHSNLSTEOILLILSLFAGHETSSVAIALAIFFLQACPKAVEELRE 341
Db 287 AGNGVGLGRLEESLSPNESADPILNLFAGNETTSTKMLFAVYFLHCKAKATOLLE 346
QY 342 LELARAKKEGSELNMDOKKMDFTQCVINFTLRNGVNFRLHRRKALDKVYKGYDIP 401
Db 347 -----DRLAGMLTWQDYKTMDFTQCVIDETRLRGLGIATMLMREKKEVSTQ 399
QY 402 GWRVLPVISAHLNDSRYDQPNLFNPMRW-----QOONNGASSSGSFSTWGNMYP 456
Db 400 GCVVVFLSAVHLDSEYKSLSFNPMRWLDPETQKRWMTSP-----FCPCFG 449
QY 457 GCPRLCAGSELAKLEMAVFIHLVLAKEELADDDOPFAFPVDFPGLPIRVSRL 511
Db 450 GGTFRCPGALRIQIALFLHYFTYTKWTQLKEDRISFPSPARLVNGFKIQCLR 504

RESULT 4
D85429
cytochrome P450 like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Apr-2001
C:Accession: D85429
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488
A:Accession: D85429
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-457 <STO>
A:Cross-references: GB:NC_001268; NID:g7270586; PIDN:CAB80304.1; GSPDB:GN00140
C:Genetics:
A:Gene: At4g36380
A:Map position: 4
C:Superfamily: Synecocystis cytochrome P450 slr0574; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:396/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 32.1%; Score 860; DB 2; Length 457;
Best Local Similarity 36.8%; Pred. No. 2.6e-48;
Matches 179; Conservative 93; Mismatches 160; Indels 54; Gaps 8;

QY 39 LPPKSGPFLGEGTIGY-----KPYATTLGDMOQHVSKYKITYRSLFGEPTIYSA 92
D 2 IPNSLGVPLVGETLNFACGYSSRPVT-----FMDKRKSLYGVFNTIGPPIIIST 55
QY 93 DAGNRFITLONEGRLEFCSEYPSRIGGILGKSMVLVVGDMRDMRSISLNFISHARLRTI 152
D 56 DAEVAKVVLQHGNTFYFAFKSTIELIGENSILSINGPHOKRLHTLIGALRSHLDR 115
QY 153 LKQVRRHTFLVDSWQNSIFSAODEAKKFTFLMAKHISMDEEETOLKKEYVT 212
D 116 ITRDIEASVNLTLASMAQLPVHODEIKMTFELIVKVLSTSPG-EDMNLKLEFEF 174
QY 213 MKGVSAFLNPGRVYHAKALOSRATILKFERKMEERLIDKEDEEVEEYKT-----E 266
D 175 IKGICIPKIPKTRLYKSLAKERLIMVKVVEERQVATTSTSPADVLLRDCD 234
QY 267 DEAEKMSKDVHKKORTDDLLGWVLKHSNSTEQILDLISLFGHETSSVAIALATF 326
D 235 SEKQSPDEPVSCK-----ITVEMLRGEETMPTAMTLAKF 270
QY 327 LQACPRAVEELREHLEIARAKKEGESELMWDDYKKMDFTQCIVINETLRGNVRFPLR 386
D 271 LSDNPVALAKLVEEMEMKRRKLEGE-EYKWTQYMSLFTQNVINETLRMANITINGYMR 329
QY 387 KALDVRKGVDIRSGKVVLYVIAVHLDNSRYDQNLFNWRMOQONNGASSSGSFS 446
D 330 KALDVEKGLVLPKGMVILASFSIVHDEDIYDNPYOFDMRMRINGSANSSIC---- 385
QY 447 TWGNVYMPFGGPRLCAGSELAKEMAVEFIHHLVLFKNMELAEDDQPPAPFVDPNGLP 506
D 386 -----FTFPGGGQRLCPLEISLKEISIFLHLVTRYSW-TAEDEIVSFPVKKRRLP 439
QY 507 IRVSRIT 512
D 440 IRVATV 445

RESULT 5
T48613
hypochemical protein F18022.190 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T48613
R:Beyan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24493

A:Accession: T48613
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-382 <BEV>
A:Cross-references: EMBL:DB
A:Experimental source: cultivar Columbia; BAC clone F18022
C:Genetics:
A:Map position: 5
A:Insertions: 31/2; 139/3; 204/1; 224/3; 253/3; 280/1; 315/3; 349/2
A>Note: F18022.190

Query Match 30.4%; Score 815; DB 2; Length 382;
Best Local Similarity 34.2%; Pred. No. 1.7e-45;
Matches 159; Conservative 87; Mismatches 127; Indels 92; Gaps 4;

QY 45 GMPFLGFTIGLYKPYATATLGDPMQOHVSKYKITYRSLFGEPTIVSADAGINRFITONE 104
D 2 GMPFGETISFFKPHRSISIGTFLQVRVSRKVKRSGNIGGKAVVSCDQELNMFITONE 61
QY 105 GRLEPCSPRSITIGILGKSMVLVVGDMRDMRSISLNFISHARLRTILLADVERHTLFV 164
D 62 GKLFTSDYPRKAMHDLGKYSLLATGELTRKLNKYNISFINLTSKPDPLCAENLSTSI 121
QY 165 LDMQONSIFSAODEAKKFTFLMAKHISMDEEETOLKKEYVTPEKMGVSAFLNP 224
D 122 LKSKNCRVEFEKREKVTFTLSVYNOLSTIKPEDPALVYLQDFLSTMGKFSIPPLP 181
QY 225 GTAVYHAKALOSRATILKFERKMEERLIDKEDEEVEEYKTDEAEKMSKDVHKKORTD 284
D 182 GTGYNAIKVRSNRNIHONAIIEEDMNNAIIEEDFLDSIISND----- 224
QY 285 DLGVLKHSNSTEQILDLISLFGHETSSVAIALATFLOACPRAVEELREHLEI 344
D 225 -----EENAI 230
QY 345 ARAKKEGESELMWDDYKKMDFTQCIVINETLRGNVRFPLRALKALDVRKGVDIRSGK 404
D 231 RAKKGDEL-LNMDYQKMEFTQCIVSEALRCGIVYTVHRAKTHDKFNEYVLPKGM 288
QY 405 VLPVIAVHLDNSRYDQNLFNWRMOQONNGASSSGSFSFTWGNVMPFGGPRLCAG 464
D 289 VEPFTAVHLDPSSLHENFEFEPNMRKTYT-----ARGGVRCVPG 329
QY 465 SELAKEMAVEFIHHLVLFKNMELAEDDQPPAPFVDPNGLP 509
D 330 GELKQLQIAFLHLVLSYRWKIKSDEMPVIAHYPVEFKRMILEI 374

RESULT 6
T07859
cytochrome P450 homolog - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 16-Feb-2001
C:Accession: T07859
R:Bishop, G.J.; Harrison, K.; Jones, J.D.
Plant Cell 8, 959-969, 1996
A:Title: The tomato Dwarf gene isolated by heterologous transposon tagging encodes th
A:Reference number: Z16181; MUID:96265705
A:Accession: T07859
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-464 <BIS>
A:Cross-references: EMBL:U54770; NID:g1421740; PIDN:AAI17070.1; PID:g1421741
A:Experimental source: strain GCR/58
C:Genetics:
A:Gene: dwarf
C:Superfamily: Synecocystis cytochrome P450 slr0574; cytochrome P450 homology
F:273-436/Domain: cytochrome P450 homology <45>

Query Match 28.4%; Score 761.5; DB 2; Length 464;
Best Local Similarity 33.4%; Pred. No. 6.2e-42;

Matches 170; Conservative 90; Mismatches 188; Indels 61; Gaps 10;

```

12 LLLPSLLSLFLILKLRN--RKTFRNLPKSGMPLFGETIGYLRKPTATTLG-DFM 68
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
5 LILSPGFCICITALLRNQYKYNQKNLPETMGMPLEGETTEFLK-----LGPSE 58
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
69 QOHVSKYKTYRNSLFGPEPTVSADAGLNRFILQNEGRLEFCSYPSNIGGILKWSMLVL 128
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
59 KNRARGFSEFKSHILGCPITVMSDELNRYIIVNEAKGLVPGYPOSMIDILKCNIAAV 118
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
129 VGDHMDRMSISLNFSLHARLRTILKDVRRHTLFVLDNQNSIFSADQAKFTFNLM 188
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
119 NSGASHYMKRALLSLISPTWIKQPLPKIDENRSHLTNM-DNKVIDIQETKMAFLSS 177
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
189 AKIMSDPGEET---EOLKEVTFPMKGVSAPLNPGTAYHAKLOSRAITLKEFERK 245
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
178 LKQI---AGIESTSLAQEMSEFNLVLTSLINLPNTVNHRRGQAKIIVNLRL 233
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
246 MERKLDIKEDQEEVEVKTEDEAKMSKDVHVKQRTDDLLGWLKHS---NLSTEQI 301
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
234 IEERR-----ASKETIOHMLGYLNNNEATRKRLTDDDM 266
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
302 LDIILSLFAGHETSSVAIALFPLQAPKAVEELREHLEIARAKKEGSELMMWDY 361
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
267 IDILITLISGTYETVSTSMAYKYLHDPKYLEELRKEHMAIREKKP--EDPIDNDY 324
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
362 KKMDFQCVINETLRIGNVVFLRKALDVRKYGDIPSGMKVLPVISAHLNDSRYDQ 421
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
325 RSMFRTRAVILLETSLATYNGVLRKTQDMETNGYIIPGKMRIVYTRRLNDPRLYPD 384
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
422 PNLFPNRRQOQNNGASSSSGSFSTWGNVYMPFGGPRLCASELAKLEMAVFIHLVL 481
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
385 PYSFNPWRMDKS-----LEHQNSELVFGGTRQCPKEKELGVAISFLHYEVT 433
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
482 KFNWELAEDDQPPFPVDFPNGLPIRVS 510
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
434 KYRWEELGDKLMPFPRVEAPNGLRIRVS 462
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 7

T04444
 cytochrome P450 - Arabidopsis thaliana
 N:Alternate names: protein T18B16.200; protein T5K18.10
 N:Contains: oxidoreductase (EC 1.-.-.-)
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 28-Jul-2000
 C:Accession: T04444; T05806
 R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Bancroft, I.; Mewes, H.W.
 submitted to the protein Sequence Database, April 1998
 A:Reference number: Z15359
 A:Accession: T04444
 A:Molecule type: DNA
 A:Residues: 1-457 <BEV>
 A:Cross-references: EMBL:AL021687
 A:Experimental source: cultivar Columbia; BAC clone T18B16
 R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voec, M.; Robben, J.; Volckaert, G.; Be
 submitted to the Protein Sequence Database, April 1998
 A:Reference number: Z15453
 A:Accession: T05806
 A:Molecule type: DNA
 A:Residues: 131-457 <BEW>
 A:Cross-references: EMBL:AL022580
 A:Experimental source: cultivar Columbia; BAC clone T5K18
 C:Genetics:
 A:Map position: 4
 A:Introns: 67/2; 173/3; 302/3; 358/1; 393/3
 A:Note: T18B16.200; T5K18.10
 C:Superfamily: Synchocystis cytochrome P450 slr0574; cytochrome P450 homology
 C:Keywords: chromoprotein; electron transfer; heme; iron; monooxygenase; oxidoreductase
 F:272-433/Domain: cytochrome P450 homology <P45>

Query Match 23.6%; Score 632; DB 2; Length 457;

Best Local Similarity 31.8%; Pred. No. 1,5e-33;
 Matches 151; Conservative 92; Mismatches 176; Indels 56; Gaps 12;

```

13 LLLPSLLSLFLILKLRNKRTRFNLPPKSGMPLFGETIGYLRKPTATTLGDFMQOHV 72
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
10 LFAGSLFLFLRCLISORRGSKKLPLPETMGMPVGET---FQLSQDP-NVFFQSKO 65
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
73 SKYKTYRNSLFGPEPTVSADAGLNRFILQNEGRLEFCSYPSNIGGILKWSMLVLGDM 132
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
66 KRYGVFKTHVLCPCVMILSSPEAKFVLVTKSHLRFKPEPASKERMLGQALFFHQGDV 125
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
133 HRDMSISLNFSLHARLRTILKDVRRHTLFVLDNQNSIFSADQAKFTFNLAHNI 192
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
126 HAKRLVLRAPMESIRN-MVPIESIAODSLRSR-EGMTWTYQEMTYTFNVALIS 183
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
193 MSMDPGEET---EOLKEVTFPMKGVSAPLNPGTAYHAKLOSRAITLKEFERMEER 249
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
184 F---GKDEVLVEDLKRCYILLEKGYNSMPVLPGLTFHRSKAKAKELSQLARLSER 239
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
250 KLDIKEDQEEVEVKTEDEAKMSKDVHVKQRTDDLLG-WVLKHSNLSTEQILLDSL 308
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
240 R-----QNGSSH-----NDLGSFMGDKRELDEQIADNITIGV 272
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
309 LFAGHETSSVAIALFPLQAPKAVEELREHLEIARAKKEGSELMMWDYKKMDFIQ 368
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
273 IFARDTASVMSWILKYLAENPNVLEAVTEQMAI-RDKEGES-LTWGDTKKMPLTS 330
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
369 CVINETLRIGNVVFLRKALDVRKYGDIPSGMKVLPVISAHLNDSRYDQPNLFNPV 428
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
331 RVIGETLRVASILSFPRFAVEDVEYEGVILIPGKMWVLPFLFRHHSADIFSNGFDP 390
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
429 RMOQOQNNGASSSSGSFSTWGNVYMPFGGPRLCASELAKLEMAVFIHLVLEKF 483
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
391 RFE-----VAPKPNTFMFGNGTHSCPNNELAKLEMSIMIHLLTRYK 432
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 8

T02739
 probable cytochrome P450 At2g29090 [imported] - Arabidopsis thaliana
 N:Alternate names: cytochrome P450 homolog T914.17
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
 C:Accession: T02739; D84692
 R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.;
 submitted to the EMBL Data Library, August 1998
 A:Description: Arabidopsis thaliana chromosome II BAC T914 genomic sequence.
 A:Reference number: Z14710
 A:Accession: T02739
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-482 <ROU>
 A:Cross-references: EMBL:AC005315; NID:g3461834; PID:g3461849
 A:Experimental source: cultivar Columbia
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,
 eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: D84692
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-482 <STO>
 A:Cross-references: GB:AE002093; NID:g3461849; PIDN:AAC33235.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g29090; T914.17
 A:Map position: 2
 A:Introns: 80/2; 187/3; 238/3; 321/3; 351/3; 413/3; 453/2
 C:Superfamily: Synchocystis cytochrome P450 slr0574; cytochrome P450 homology
 F:291-453/Domain: cytochrome P450 homology <P45>

Query Match 22.8%; Score 611; DB 2; Length 482;

Best Local Similarity 29.8%; Pred. No. 3, 5e-32;
Matches 153; Conservative 101; Mismatches 192; Indels 68; Gaps 13;

```
QY 16 PSLSLLELLILK-----RRNRKTRFNLPPKSGMPLEGITIGYLKPYATATLGDQM 69
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 20 PALITLTIVVVVVLLFKWMLHWKEQRLRLPPSGMLPTIGER--LRLYTEMP-NSFFA 75
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 70 QHVSRYGKIKRNLGEPPIVSADAGLNFLIONERLEPCSPRISIGILGKWSMLVLV 129
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 76 TRONRYGIFGFKHIGCCPMVMSPEARMVLYSKAHLLKPYPPSKERMIGEGALFEHQ 135
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 130 GDMHDM-RSISINFLSHARLRTILLKQVERHTLFLYDSMQMSIFSADQAKKFTFNL 188
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 136 GPHSTLRKLVOSFSPSLRPTV--SHIELVLYQLTSMTSOKSINTLEYMKRYADVA 193
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 189 AKHINSMDPGEE---TEQLKREYVFMKGVVSAPINLPGTAYAKALQSRATILKIER 245
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 194 ---ISAGDKKEPTITIDIKILYQLRERGYNMPLDITGLTFHKSKAKAIELEELRY 250
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 246 MEERLIDKEDEDEBEVKTEDAEKSKSDHVRKQRTDDLGLGVL---KHSNISTEQ 300
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 251 IEKRRENGREE-----GGLLVLLGAKQDKRNGLSDSQ 283
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 301 ILDLISLFLNCHETSSVAIALAIFFLQCPKAVELELREHLEI-ARAKELGESELNWD 359
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 284 IADNLTIGVFAATDTTASVLTMLKYLHDHPNLQLEVSREQFSIRQIKKE--NRRTSWE 341
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 360 DYKKNDPQCIVNETLRLGNVVRFLRKALKDVRKGYDPSQMKPLPVSAVHLDSRY 419
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 342 DTRKMLPLTRVIGELTRASVLSFTFRNAVQDVEYGYLIRPKMKVLYPLRRIRHNSFF 401
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 420 DQPNLFNPMRWQOQNNNGASSGSGSFSTWGNMMPGCGPRCLCAGSELAKLENAVFIHL 479
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 402 PDPEKEDSRRFE-----VAPRPYTYMPGNGVHSCPSSELAKLEMLILLHL 448
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 480 VLKFWELAEDDQDPFAF-PFVDFPNCGLPIRVSR 512
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 449 TTSEFRWEVIGDEGIQYGPFPVPRKGLPIRVPTI 482
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 9
C71417
Cytochrome P450 d13695c - Arabidopsis thaliana
N:Contains: oxidoreductase (EC 1.-.-.)
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: columbia
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #ext_change 16-Jun-2000
C:Accession: C71417
R:Bayan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk,
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel,
A.; vanagh, T.; Hempel, S.; Kotter, P.; Ertlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reichman, S.; Ans
C.; Chaitany, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha
A:Reference number: A71400; MUID:98121113
A:Accession: C71417
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-487 <BEV>
A:Cross-references: GB:297338; NID:92244870; PIDN:CABI0309.1; PID:92244888
C:Gene: d13695c
A:Map position: 4COP9-4G3845
C:Superfamily: Synchocystis cytochrome P450 slr0574; cytochrome P450 homolog
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
F:433/Binding site: heme iron (Cys) (axial ligand) #status predicted
```

Query Match 21.7%; Score 581; DB 1; Length 487;
Best Local Similarity 27.3%; Pred. No. 3, 1e-30;
Matches 143; Conservative 100; Mismatches 205; Indels 76; Gaps 11;

[illegible]

Db 72 EDVNPRIKIKNGIKFKIRINGSPTIIVNGAEANRLILSNEFSLVSSWNPSSVOLMGNC 131
 Qy 125 MLVLGDMHDMRSISLNFSLHARLRTLLKDVHRTLFVLDWMQONSIFSADAEAKFT 184
 Db 132 IMAQGEKHVRLRGIVANSLSYIGLESILPKLCPTVKFHNETEMRGKREISLYSACKLT 191
 Qy 185 ENLAKH--MSMDGEETDLKKEYTFMKGVASAPLNPRTAYAKALOSRATILKFI 242
 Db 192 FTVVEECYLGKIVEIGMLEV-----FERVLEGFALFVEPEPCSFARAKKARIEITFL 245
 Qy 243 ERKMERKLDIKEEDOEDEEVEKTEDEAMSKSDVHRKQRTDDLLGWLKHNSNSTEOL 302
 Db 246 VGRKREKREKREKGA-----KPNITLSRLVEELIKGV-----ITEEV 287
 Qy 303 DLISLFAGHETSSVALAIAIFLQACPKAVEELREHELEIARAKKEGESE-LMMDY 361
 Db 288 DNVLLVFAADHTSYAASMTFMKLAQHPTCRDITLQENHAQI---KANKGEXELTVEDV 344
 Qy 362 KKMFTQCVINETRLGNVAVFLHRKALKDVRKGYDIPSGWKVLPVISAHVLDNSRYDQ 421
 Db 345 KKMYSQVQVRETMKLSPPRIGSFRKAVADIDYGTYIPKGMKILMTYTGTHYNEPEIFQD 404
 Qy 422 PNLFPNPMQOONNGASSSGSFTWGNMYPFGGGRCLAGSELAKLEMAVFIHLVL 481
 Db 405 PMSDPTREFDKPIQAYT-----YLPRGGGRCLAGHQLAKISTLVPMHEVYT 451
 Qy 482 KFMNELAEDDQFAFPVYDPFN-GLPIRVS 510
 Db 452 GFDMSLVYPRDETISMDFPPLSGMPIKIS 481

RESULT 11

H86185
 hypothetical protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: H86185
 R:Phenologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 ansen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: H86185
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-490 <STO>
 A:Cross-references: GB:AE005172; NID:92388581; PIDN:AAB71462.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: Synchocystis cytochrome P450 slr0574; cytochrome P450 homology

Query Match 21.6%; Score 579.5; DB 2; Length 490;
 Best Local Similarity 28.6%; Pred. No. 4e-30;
 Matches 133; Conservative 107; Mismatches 192; Indels 33; Gaps 9;

Qy 34 KTRNLPKSGMPLGFTIGLKPYATTLGDFMQOHVSKYK--IYRSNLFSEPTIVS 91
 Db 40 ENRHYLPKPGDGLPFGIMLSLRAFKTSDPSFTRLIKRYGPKGIYKAMHFNPSITV 99
 Qy 92 ADAGLNRFILONEGRLEPCSPRSIGILGKSMVLVLDGHRDMRSISLNFSLHARLT 151
 Db 100 TTSOTCRVLTDD--AFKPGWPTSMELTGKSKSVGISFEHAKRRLRTLAAPVNGHEALS 158
 Qy 152 ILKDVHRTLFVLDWMQONSIFSADAEAKFTFNLMAKHMSMDPGEETEOLKKEYVT 211
 Db 159 TYIYEIENVTIVLDKMTKMGEEFLTHLRKLTFRIT--MYFLSSESNVADALEREYTA 217

Qy 212 FMKGVASAPLNPRTAYAKALOSRATILKFIERKMERKLDIKEEDOEDEEVEKTEDEAM 271
 Db 218 LNVGVAAMAVNIPGAVYHARLAKARTLVANQSVTER-----NQKQVILNSKKMVL 271
 Qy 272 SKSDHVRKQRTDDLLGWLKHNSNSTEQLDILSLLFAGHETSSVALAIAIFLQACP 331
 Db 272 DNLNVRK-----DED-----GKTLDEEITIDVILMYLMAHESSGHTIMATVFLQENP 320
 Qy 332 KAVELEREHELEIARAKKEGESELMMDYKKMFTQCVINETRLGNVAVFLHRKALKD 391
 Db 321 EVLQAKAEQMLKSRP--GQKLSLKETKMEFLSVYDETRIVITFSLTAREAKTD 379
 Qy 392 VRYGYDIPSGWKVLPVISAHVLDNSRYDQNLFPNPMQOONNGASSSGSFTWGNM 451
 Db 380 VEMNGYLLPKGMKVLTFMRDYNHIDPEVPPDRKFPDARW---DNG-----FVPRAGA 428
 Qy 452 YMPFGGRCLAGSELAKLEMAVFIHLVLKFMNELAEDDQFAF 496
 Db 429 FLPRGAGSHLCPGNDLAKLETISFLHFLKTYQVRSNPECPVM 473

RESULT 12

T48973
 cytochrome P450-like protein - Arabidopsis thaliana
 N:Alternate names: protein F14D17.40
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
 C:Accession: T48973
 R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: 225008
 A:Accession: T48973
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-455 <JDR>
 A:Cross-references: EMBL:AL353992; GSPDB:GN00061; ATSP:F14D17.40
 A:Experimental source: cultivar Columbia; Bac clone F14D17
 C:Genetics:
 A:Gene: ATSP:F14D17.40
 A:Map position: 3
 A:Insertions: 70/2; 178/3; 229/3; 312/3; 343/3; 346/1; 381/3
 C:Superfamily: Synchocystis cytochrome P450 slr0574; cytochrome P450 homology

Query Match 21.5%; Score 577; DB 2; Length 455;
 Best Local Similarity 28.9%; Pred. No. 5.2e-30;
 Matches 145; Conservative 84; Mismatches 173; Indels 100; Gaps 10;

Qy 32 NKRTRNLPKSGMPLGFTIGLKPYATTLGDFMQOHVSKYKGIYRSNLFSEPTIVS 91
 Db 28 NPKSNGKLPKSGMPLGFTIGLDFKPYGEYEISPYLKKMLKRLPFRITNLTGVKTVVS 87
 Qy 92 ADAGLNRFILONEGRLEPCSPRSIGILGKSMVLVLDGHRDMRSISLNFSLHARLT 151
 Db 88 TDKVYHMEILKQENKSFILSTPDGIMKPLKDSLFKIGNHKKIKQITLLLSSEGILK 147
 Qy 152 ILKDVHRTLFVLDWMQONSIFSADAEAKFTFNLMAKHMSMDPGEETEOLKKEYVT 211
 Db 148 KILKDMRVRTREHLSNAKTRGLDVKDAVSKLILAHILTPKMS-----NLKPTQA 198
 Qy 212 FMKV-----VASPLNPRTAYAKALOSRATILK--FIERKMERKLDIKE 255
 Db 199 KLMGIFKAFTFDWTSTYLSAGGLVNTLM--ACRGMREIKIDYIMRKTSSEKY--- 252
 Qy 256 EDGEEVEKTEDEAMSKSDVHRKQRTDDLLGWLKHNSN-----LSTEQILDILSLF 310
 Db 253 -----DPLNLAIESEKAAGELNDENAITTLFLTSC 284
 Qy 311 AGHETSSVALAIAIFLQACPKAVEELREHELEIARAKKEGESELMMDYK--KMFQOC 369
 Db 285 VTQDTSKALCLAVKFLLENKQVLAEKKEH--EVLISREDEKGGVYMEVYRHNMFTFN- 342

T02263
 cytochrome P450 DWARF3 - maize
 N:Contains: oxidoreductase (EC 1.-.-.-)
 C:Species: Zea mays (maize)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T02263
 R:Winkler, R.G.; Helentjaris, T.
 Plant Cell 7, 1307-1317, 1995
 A:Title: The maize dwarf3 gene encodes a cytochrome P450-mediated early step in gibberellin biosynthesis
 A:Reference number: Z14648; MUID:96004534
 A:Accession: T02263
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-519 <WIN>
 A:Cross-references: EMBL:U32579; NID:9987266; PIDN:AAC49067.1; PID:9987267
 A:Experimental source: strain B73
 C:Genetics:
 A:Gene: dwarf3
 C:Function:
 A:Description: Involved in an early step in gibberellin biosynthesis
 A:Pathway: gibberellin biosynthesis
 C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology
 C:Keywords: oxidoreductase
 F:325-488/Domain: cytochrome P450 homology <P45>

Query Match 18.3%; Score 491.5; DB 1; Length 519;
 Best Local Similarity 28.2%; Pred. No. 2,1e-24;
 Matches 129; Conservative 91; Mismatches 189; Indels 49; Gaps 11;

QY 36 RFNLPPGKSGMPLIGETIGYLYKPYTATLGDPMQOHVSKYK--IYNSLFGPEPTI-VSA 92
 Db 69 RARLPPEMGWPLVGGWMAFRAKSGKPDAFIAFVRFRGTGVYRSFMSSPTVLT 128
 QY 93 DAGLNRFILONEGRLEPCSYRSTGILGKWSMLVLVGDMDRSTSLNLSHARLTI 152
 Db 129 AEGCKQVIMDDA--FVTGMKATVALGPRSEYAMPYDEHRLRLKLTAPDINGFDALTG 186
 QY 153 LKLDVERHTLEFLDSW-QQNSIFSADDEAKKFTENLAKHIMSMDPGEETEQLKKEYVT 211
 Db 187 YLPFIDRTVTSILRAMADHGGVSEFTELRRMTEKTIYQ-IFLGADQATRALERSYTE 245
 QY 212 FMKGVNSAPLNLPGTAHKKALQSRATILKTERKMEERKLDIKE-----EDQEEVEYKT 265
 Db 246 LNYGMRAMAINLPGFAYRGALRRARRLVAVLQGVLDERRARAKAVSGGVDMDRLIEA 305
 QY 266 EDEAEMSKSDHVRKQRTDDDLGWLKHSNLSSTQILDILSLFAGHETSSVALALAI 325
 Db 306 QDE-----RGRHDD-----EITDVLVMTLNAGHSSGHTMTATV 342
 QY 326 FLQACPRAVEELREHELEIARAKKEGSELNMDYKMDFTQCVINETLRLGNVRF 385
 Db 343 FLQENPDMFAKAEQALMRSIPS-SORGLTLNDFRMEYLSQVIDETLALVNISFVSF 401
 QY 386 RKALKDVRKYCYDLPISGKYLPIVISAHLNLSRYDQRLFPNPMWQOONNGASSSGSGSF 445
 Db 402 RQATRDVFNVGYYLIPKGMKQVLWYRSVAMDPQVYDPPTKFDPSRME----GHSRAGTF 456
 QY 446 STMGNNYMPGGGRLCAGSELAKLEMAVFTNHLVKF 483
 Db 457 -----LARGLGARLCRGNDLAKLEISVFLNHFLLGY 487

Search completed: September 28, 2001, 18:45:05
 Job time: 95 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2001, 18:43:30 ; Search time 14.33 Seconds
(without alignments)
737.114 Million cell updates/sec

Title: US-09-502-426-2

Perfect score: 2681

Sequence: 1 MFTTEHHTLLPLLLPLSLLS.....FAFPVDFPENGLPFRYSRL 513

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 segs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	978.5	36.5	472	2	US-08-622-166A-2
2	978.5	36.5	472	2	US-08-622-166A-4
3	503	18.8	492	3	US-08-724-466B-2
4	439	16.4	497	3	US-08-724-466B-4
5	289	10.8	511	4	US-08-991-677-4
6	279.5	10.4	513	3	US-08-948-564-6
7	264	9.8	520	2	US-09-091-432-2
8	263.5	9.8	504	1	US-08-457-274A-25
9	263.5	9.8	504	5	PCT-US95-05758-25
10	257	9.6	476	1	US-08-313-075A-30
11	256.5	9.6	490	1	US-08-201-118-7
12	256.5	9.6	490	2	US-08-238-821B-7
13	256.5	9.6	490	5	PCT-US95-05744-7
14	245	9.1	495	1	US-08-532-065B-2
15	244.5	9.1	576	3	US-08-948-564-16
16	241	9.0	510	3	US-08-606-505B-66
17	241	9.0	510	4	US-09-616-990-66
18	240.5	9.0	509	3	US-08-948-564-18
19	239	8.9	496	1	US-08-313-075A-50
20	236	8.8	521	1	US-08-948-564-14
21	234.5	8.7	490	1	US-08-201-118-1
22	234.5	8.7	490	2	US-08-238-821B-1
23	234.5	8.7	490	5	PCT-US95-05744-1
24	233.5	8.7	508	4	US-08-991-677-2
25	231	8.6	506	1	US-08-313-075A-38
26	231	8.6	506	3	US-08-606-505B-65
27	231	8.6	506	4	US-09-616-990-65

28	223.5	8.3	496	4	US-09-292-768-64	Sequence 64, Appl
29	223	8.3	490	1	US-08-201-118-3	Sequence 3, Appl
30	223	8.3	490	2	US-08-238-821B-3	Sequence 3, Appl
31	223	8.3	490	5	PCT-US95-05744-3	Sequence 9, Appl
32	222	8.3	490	1	US-08-201-118-9	Sequence 9, Appl
33	222	8.3	490	2	US-08-238-821B-9	Sequence 9, Appl
34	222	8.3	490	5	PCT-US95-05744-9	Sequence 9, Appl
35	222	8.3	516	3	US-08-948-564-12	Sequence 12, Appl
36	221.5	8.3	500	2	US-08-314-601-2	Sequence 2, Appl
37	221.5	8.3	500	5	PCT-US95-13051-2	Sequence 2, Appl
38	218	8.1	510	3	US-08-948-564-4	Sequence 4, Appl
39	215	8.0	500	3	US-08-881-784-9	Sequence 9, Appl
40	214.5	8.0	496	3	US-08-881-784-1	Sequence 1, Appl
41	214.5	8.0	496	4	US-09-292-768-2	Sequence 2, Appl
42	214.5	8.0	496	4	US-09-292-768-66	Sequence 66, Appl
43	214	8.0	490	1	US-08-201-118-13	Sequence 13, Appl
44	214	8.0	490	2	US-08-238-821B-13	Sequence 13, Appl
45	214	8.0	490	5	PCT-US95-05744-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-622-166A-2
Sequence 2, Application US/08622166A
Patent No. 5952545
GENERAL INFORMATION:
APPLICANT: KONZ, CSABA
APPLICANT: MATUR, JAIDEEP
APPLICANT: SZEKERES, MIKLOS
APPLICANT: ALTMANN, THOMAS
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/622,166A
FILING DATE: 27-MAR-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 0147-0153P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-622-166A-2
Query Match 36.5%, Score 978.5, DB 2, Length 472:
Best Local Similarity 41.9%, Pred. No. 21e-89,
Matches 210, Conservative 79, Mismatches 171, Indels 41, Gaps 8,
QY 12 LLLPLSLILFLILKKRRNRKTRFNLPPGKSGPFLGFTIGYIKPYATVATLGDPMOQH 71

```

Db 7 LLLLSIAAGL---LLRRTYRMGLPPGSLPLIGETFOIGAKTENPEPFI DER 63
QY 72 VSKYGIYRSNLFGEPTIVSADAGINFTIIONEGRLFECSPRSIGGILGKMSLVYGD 131
Db 64 VARYGVFMTHLFGEPTIFSADPETNRFVLQNECKLFECSPASICNLGKHSLLMKGS 123
QY 132 MHRDMRSISLNFSLHARKRTLLKDVREHTLFVLDMSQONSIFSADPAKKTFTNLMAKH 191
Db 124 LHKRMSLTMSPANSIILKDHMLDRLVFNLDSSRVL--MEAKKITFELTVKQ 181
QY 192 IMSMDPEEETBOLKKEYVTFMKGVSAPLNLPCTAYHKALOSATILKTERKMEERKL 251
Db 182 LMSDPG-EMSESLRKEYLVIEGFSPLPLFSTYRKAIQAR-----RKVAEALT 232
QY 252 DIKEDEDEEVEKTEDAEKSDHVRKORTDOLGWLKHSNLSTEOILDLISLFA 311
Db 233 VVWKKRREEBEGAE-----RKDMLAALLAADGFSDEIYDFVALLVA 278
QY 312 GHETSSVAIALAIFLOACRAVELEBEHLEIRAKKEGESELDNDYKKMDFTCVI 371
Db 279 GYETSTIMTLAVKFLFETPLALQLEKEHEKIRAKMSD--SYSLWSDYKSMPEFCV 336
QY 372 NETLRGNVYRFLHRAKLDVRYKGYDIPSGKVLPIYSAVHLDNSRYDQNLFPNRMQ 431
Db 337 NETLRANITIGVFRRAITVEIKGYIKPGWKVSSFRVHLDNPFKARFTNPMRWQ 396
QY 432 QONNGASSSGSFSTGNN-YMPFGGPRICAGSELAKLEMAVFIHLVLKFMWELAD 490
Db 397 -----SNSVTTGPSNVFTFPGGPRICPEYELARVALSVFLRLVYGFWSVPAEQ 446
QY 491 DQPAFPFVDEPNCGLPIVSR 511
Db 447 DKLVFFPTTRTKRYPIFKR 467

RESULT 2
US-08-622-166A-4
; Sequence 4, Application US/08622166A
; Patent No. 5952545
; GENERAL INFORMATION:
; APPLICANT: KONCZ, CSABA
; APPLICANT: MATHUR, JAIDEEP
; APPLICANT: SZEKERES, MIKLOS
; APPLICANT: ALTMANN, THOMAS
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolaach & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/622,166A
; FILING DATE: 27-MAR-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 0147-0153P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEEX: 248345
; INFORMATION FOR SEQ ID NO: 4:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-622-166A-4

Query Match 36.5%; Score 978.5; DB 2; Length 472;
Best Local Similarity 41.9%; Pred. No. 2.1e-69;
Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;

QY 12 LLLPSLSLLFIILKRRNRKTRFNLPRGKSSWPRFLGFTIGLYATTLGDEMOQH 71
Db 7 LLLLSIAAGL---LLRRTYRMGLPPGSLPLIGETFOIGAKTENPEPFI DER 63
QY 72 VSKYGIYRSNLFGEPTIVSADAGINFTIIONEGRLFECSPRSIGGILGKMSLVYGD 131
Db 64 VARYGVFMTHLFGEPTIFSADPETNRFVLQNECKLFECSPASICNLGKHSLLMKGS 123
QY 132 MHRDMRSISLNFSLHARKRTLLKDVREHTLFVLDMSQONSIFSADPAKKTFTNLMAKH 191
Db 124 LHKRMSLTMSPANSIILKDHMLDRLVFNLDSSRVL--MEAKKITFELTVKQ 181
QY 192 IMSMDPEEETBOLKKEYVTFMKGVSAPLNLPCTAYHKALOSATILKTERKMEERKL 251
Db 182 LMSDPG-EMSESLRKEYLVIEGFSPLPLFSTYRKAIQAR-----RKVAEALT 232
QY 252 DIKEDEDEEVEKTEDAEKSDHVRKORTDOLGWLKHSNLSTEOILDLISLFA 311
Db 233 VVWKKRREEBEGAE-----RKDMLAALLAADGFSDEIYDFVALLVA 278
QY 312 GHETSSVAIALAIFLOACRAVELEBEHLEIRAKKEGESELDNDYKKMDFTCVI 371
Db 279 GYETSTIMTLAVKFLFETPLALQLEKEHEKIRAKMSD--SYSLWSDYKSMPEFCV 336
QY 372 NETLRGNVYRFLHRAKLDVRYKGYDIPSGKVLPIYSAVHLDNSRYDQNLFPNRMQ 431
Db 337 NETLRANITIGVFRRAITVEIKGYIKPGWKVSSFRVHLDNPFKARFTNPMRWQ 396
QY 432 QONNGASSSGSFSTGNN-YMPFGGPRICAGSELAKLEMAVFIHLVLKFMWELAD 490
Db 397 -----SNSVTTGPSNVFTFPGGPRICPEYELARVALSVFLRLVYGFWSVPAEQ 446
QY 491 DQPAFPFVDEPNCGLPIVSR 511
Db 447 DKLVFFPTTRTKRYPIFKR 467

RESULT 3
US-08-724-466B-2
; Sequence 2, Application US/08724466B
; Patent No. 6063606
; GENERAL INFORMATION:
; APPLICANT: Petkovich, P. Martin, White, Jay A.,
; APPLICANT: Beckett, Barbara R., Jones, Glenville
; TITLE OF INVENTION: Retinoid Metabolizing Protein
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Blake, Cassels & Graydon
; STREET: Box 25, Commerce Court West
; CITY: Toronto
; ZIP: M5L 1A9
; COUNTRY: Canada
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
; OPERATING SYSTEM: MS-DOS 5.1
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,466B
; FILING DATE: October 1, 1996
; PRIOR APPLICATION DATA:

```


APPLICATION NUMBER: 08/667,546
FILING DATE: June 21, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunt, John C.
REGISTRATION NUMBER: 36,424
REFERENCE/DOCKET NUMBER: 50767/00004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 863-4344
TELEFAX: (416) 863-2653
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 492 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-724-466B-2

Query Match 18.8%; Score 503; DB 3; Length 492;
Best Local Similarity 27.6%; Pred. No. 8.4e-42;
Matches 144; Conservative 95; Mismatches 216; Indels 66; Gaps 14;

7 HTLPRLPLPSLLSLLEF-----ILKRRNRKTRFNPFGSGWPFGLGTIGYLP 58
119 ILGKSMVLVVDGMDHRD-----MRSISLNFSLSHARLRTLLKDVERTHTLVLDSDW-QQNS 172
119 ILGSGTSLNVHGVQKNNKKKAIMRAFSRDLH-----YIVIQEVKSAIOENLQKNS 172
173 IFSADDEAKFFPNMAKIHMSDPEEET--EQLKEYVTMKGVASPLMPTGTAHK 230
173 CVLVYPRMKKIFRIAMRLLGFEEPOLITDEQLVEAFEEEMIKNLSPIIDVPSGLYR 232
231 ALQSRATILKFERMKEERKLDIKEDDEEVEKTEDEAEKMSDHYVRKQRTDDDLGAV 290
233 GLRAR---NFIHSTKEEIRKKIODDNEENOKYKDALQLL-----TENSRSDE----- 279
291 LKHSNLSTEQIIDLILSLFAGHETSSVALAIFFLQACPRAVEELREHLEIARAKE 350
280 ----PFSIQAMKEATFELFGHETTASTATSLVWELGINTVEYQKRE-----VQEKVE 331
351 LG-----ESELNDDYKKMDFTQCVINETRLGNVRFELRKALKDVRKYGYDIPSGWYL 406
332 MGMTYPRGKLSMELLDOLKYTGCIKETLRINPVPVGGCFRVALKTFELNGYQIIPKGMVY 391
407 PYISAVHLDNSRYDQPNLFNPRMOQONNGASSSGSSTGNNYMPGGCPRLCAGSE 466
392 YSICDTHDAVDFPNKEEFQPERFM-----SKGLEDSRF--NIIPEGGSRMCVGE 442
467 LAKLEMAVEFIHHLVLFKNWELAEDDO-----PFAFPEVDFP 502
443 FAKVLKLTIVLLELTHQCNWILSNGPTMTKTGPTIYVVDMLP 483

RESULT 4
US-08-724-466B-4
Sequence 4, Application US/08724466B
Patent No. 6063606
GENERAL INFORMATION:
APPLICANT: Belkovich, P. Martin, White, Jay A.,
APPLICANT: Beckett, Barbara R., Jones, Glenville
TITLE OF INVENTION: Retinoid Metabolizing Protein
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Blake, Cassels & Graydon
STREET: Box 25, Commerce Court West
CITY: Toronto
ZIP: M5L 1A9

COUNTRY: Canada
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 MB storage
COMPUTER: COMPAQ, IBM PC compatible
OPERATING SYSTEM: MS-DOS 5.1
SOFTWARE: WORD PERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,466B
FILING DATE: October 1, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,546
FILING DATE: June 21, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunt, John C.
REGISTRATION NUMBER: 36,424
REFERENCE/DOCKET NUMBER: 50767/00004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 863-4344
TELEFAX: (416) 863-2653
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-724-466B-4

Query Match 16.4%; Score 439; DB 3; Length 497;
Best Local Similarity 27.0%; Pred. No. 2.2e-35;
Matches 142; Conservative 85; Mismatches 226; Indels 72; Gaps 15;

10 LPPLLLPSL-----LSLLFLILLL-----RRNRKTRFNPFGSGWPFGLGTIGYLP 57
3 LPPLASALCTVLPPLFLAIIKMDLVCSGRDSCALPLPPTGMPFGEYLP----- 58
58 PYATATLIGPMOONHVKYKTIYRSNLFGEPTIVSADAGLNRFLLQNGHLEFECSTYRSIG 117
59 -QWVLRKREFLQMKRRKRYFIYKTHLFGRTYRVMGADVRRLLDGLDVLVSHMPASVR 117
118 GLIGKSMVLVVDGMDHRDMSISLNFSLSHARLR--TILKDVERTHTLVLDSDWQ--NS 172
118 TILGSGCLSNLHDSHQRKRYIMRAFSREALCYVPVITEVGS--LEQWLSCGR 173
173 IFSADDEAKFFPNMAKIHMSDPEEET--GEETPOLKEYVTMKGVASPLMPTGTAHK 227
174 GLVYPRVPRMKKIFRIAMRLLGFEEPOLITDEQLVEAFEEEMIKNLSPIIDVPSGLYR 232
228 YHKLQSRATILKFERMKEERKLDIKEDD-----OEEVEKTEDEAEKMSDHYVRKQRT 282
233 LYRGAKARLLIARIQONIRAKICGLRASEAGCGKDALQLLIEHSE-----RGERL 285
283 DDGLGLVNLKHSNLSTEQIIDLILSLFAGHETSSVALAIFFLQACPRAVEELREHLE 342
286 DMQ-----ALKOS--STE-----LLFGHETTASATSLTYLGLVPHVLOKVEELK 331
343 ELARAKKEIGESLMDYKKMDFTQCVINETRLGNVRFELRKALKDVRKYGYDIPSG 402
332 SKGLCKSNQDKLMDLLELQKYTGCIKETLRINPVPVGGCFRVALKTFELNGYQIIPK 391
403 WKVLPYISAVHLDNSRYDQPNLFNPRMOQONNGASSSGSSTGNNYMPGGCPRLCAGSE 462
392 WNYIYISICDTHDAVDFPNKEEFNPRFSAPHEDSRFS-----FIPPGGLRSG 442
463 AGSELAKEMAVEFIHHLVLFKNWELAE-----DDQFAPPEVDFP 502
443 VKGEFAKILKLTIVLLELTHQCNWILSNGPTMTKTGPTIYVVDMLP 487

RESULT 5
US-08-724-466B-4
Sequence 4, Application US/08991677A
Patent No. 6252135

```

: GENERAL INFORMATION:
: APPLICANT: Chiang, Vincent L
: APPLICANT: Carraway, Daniel T
: APPLICANT: Smeltzer, Richard H
: TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
: FILE REFERENCE: 50617
: CURRENT APPLICATION NUMBER: US/08/991,677A
: CURRENT FILING DATE: 1997-12-16
: EARLIER APPLICATION NUMBER: US 60/033,381
: EARLIER FILING DATE: 1996-12-16
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 4
: LENGTH: 511
: TYPE: prt
: ORGANISM: Liquidambar styraciflua
: US-08-991-677-4

Query Match          10.8%   Score 289;   DB 4;   Length 511;
Best Local Similarity 22.0%   Pred NO. 2.3e-20;
Matches 121; Conservative 95; Mismatches 196;   Indels 138;   Gaps 21;

```

```

RESULT 7
US-09-091-432-2
; Sequence 2, Application US/09091432
; Patent No. 5981837
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clint
; TITLE OF INVENTION: A Method For Regulation Of Plant Lignin Composition
; FILE REFERENCE: 7024-325
; CURRENT APPLICATION NUMBER: US/09/091,432
; EARLIER FILING DATE: 1998-06-18
; EARLIER APPLICATION NUMBER: PCT/US96/20094
; EARLIER FILING DATE: 1996-12-19
; EARLIER APPLICATION NUMBER: US 60/009,119
; EARLIER FILING DATE: 1995-12-22
; EARLIER APPLICATION NUMBER: US 60/013,388
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Microsoft Word 2.0C
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Sequence is deduced from DNA sequence of SEQ ID NO:1
; Patent No. 5981837
US-09-091-432-2

```

```

Query Match          9.8%; Score 264; DB 2; Length 520;
Best Local Similarity 20.3%; Pred. No. 7,6e-18;
Matches 109; Conservative 103; Mismatches 208; Indels 118; Gaps 18;

```

```

QY 11 PLLDLSLLFLILKRRNKRTPGSGWPFLETTGYLKPATATLGDPMQ 70
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 15 PTLSTLVVSLFISFISFIRRRPP---YPPGPRGPIIGNML-MMDQLHRLANDL--- 67
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 71 HVSXGKITYRSLNFGFPTIVSADAGLNRTLLQNEGRLE-----CSYRSISGIL 120
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 68 -ARKYGLCHLRMGLHMTAVSSPEVAROVLQVDSFNSRPATIAISYLTDRAD--- 122
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 121 GKMSMLVLVGMHMRDRSIS-LNFLSHARLTILKDVHERHTLFLVDSMOQNSIFSAD 179
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 123 ---MAFAHNGPFRQMRKVCVMKVFPSRKR-----AESMA-----SYRDE 158
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 180 AKKFTFNL-----MAKHNSM-----DPGEETEQLKKEVTFEMKV 216
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 159 VDKMVRSVSCNGKRVNGBQIPALTRNITYRAAFSGACEKGDEFIRLQEF-SKLFGA 217
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 217 VSAFLMLPGTAY-----HKALQSRATILKFERKMEERKLDIKEEDQEEVKTDE-DE 268
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 218 FNVADPIYFVGWIDPQGINRKRLKARNDLDGFLD-----DIDEHMKKKKNQNAVD 269
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 269 AEMSKSDHYAKQRTDDLLGLWVLKHSNLSTE-----QLDILTLISLFAGHET 315
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 270 GDVVDIDM-----DDLAFYSEAKLYSETFDLQNSIKTLTNDNKAITIMDMVFGTET 323
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 316 SSVAIALAIFELQACRAVEELREHELETARAKKEGESELANDDYKKMDFTOCVINETL 375
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 324 VASAIEMALTELRLSPEDLTKRYQQLAEVGLDRVVEES-----DIEKTYLLACTIKETL 378
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 376 RGNVVRFLHRRKALKDVRKYGYDIPSGWKVLPYISAVHLDNSRYDQNLFPNPRMOQNN 435
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 379 RMPPTPLLHETADSTISDGFPIPKSKRWMAFAIGRPTJMTDPTDRPSFLE--- 435
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 436 GASSSGSGSFSTWGNMVMPPGGGPRLCAGSELAKLEMAVFIHLVLKFMNELAEDDP 493
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 436 ----PGVPDRKSGSNFEIPIPGSGRRSGPQMGDLGYALDLAVAHILHCFITKLDGKMP 489
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RESULT 8
US-08-457-274A-25

```

```

; Sequence 25, Application US/08457274A
; Patent No. 5734086
; GENERAL INFORMATION:
; APPLICANT: Scott, Jeffrey G.
; APPLICANT: Tomita, Takashi
; TITLE OF INVENTION: Cytochrome P450Lpr Gene and Its Uses
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,274A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1600
; TELEFAX: 716-263-1304
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Rat
; STRAIN:
; DEVELOPMENTAL STAGE: Adult
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
US-08-457-274A-25

```

```

Query Match          9.8%; Score 263.5; DB 1; Length 504;
Best Local Similarity 24.1%; Pred. No. 8.1e-18;
Matches 132; Conservative 89; Mismatches 213; Indels 113; Gaps 25;

```

```

QY 9 LPLDLLPSLLSLFLILKRRNKRTPGSGWPFLETTGYLKPATATL 64
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 3 LLSALTLFETVLLAVLVLLYGFGRTHGLFKQGIQGRPRDPFGVLYUY-----MGL 57
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 65 GDMQOQVSKYKGIYKSNLF-GE-PTIVSADAGLNRTLLQNEGRLEPC-----SYRSG- 117
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 58 WKFDVDECHKRYKTIW-GLFDGMPLEPAITDTEIKNVL-----VKCFSEVFNRRDFGP 110
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 118 -GILGKMSMLVLVGMHMRDRSISLNFSLSHARLTILKDVHERHTLFLVDSMOQNS---- 172
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 111 VGIMGRAVSAKDEMKRYALSLPTTS-GRLE-MFLIEQYGLDVLVYLKQEAETGK 168
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 173 -----IFSA--ODEAKKFTFNLMAKHNS-MDGEETEQLK--EYTFMKGVVASPL 221
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 169 PYTMKRVFGAYSMDVITSTSGVNVDSLNPDKDPFEKTKLLRDFDPLFLSVLFP 228
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 222 NLPGTAVHKL-----QGRATILKFERKMEERKLDIKEEDQEEVKTDEDAEKSKS 274
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 229 LTP---IYEMLNICMPKSDIEFFKRVYR-MKETRL----- 261
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 275 DHVRKQRTDDLLGLWVL-----KHSNLSTEQLDILTLISLFAGHETSSVAIALAI 324
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

Db 262 DSVQKHRV--DFLOLMNAHNDKSKDESHETALSDMEITAOISIFIFAGYEPTSSITLSFVL 319
QY 325 FFLQACPRAVELEHREHLEIARAKKEGESELMNDYKMKDFTQCVINETRLGNAVREL 384
Db 320 HSLATHPTQKKLOE---EIDRALP--NKAPPTYDTVEMEMELDMVLEITRLPIGRL 374
QY 385 HRKALKDVRKGYDIPSGWKVLPVISAVALDNSRYDQNLFPNPMWMOOONNGASSSGSGS 444
Db 375 ERVCKKDYEVGFMFKGSVMIPISYALHRDPQHWPPEEPERPERFSKENGSDP---- 430
QY 445 FSTGNNNTMPFGGPRLCAGSELAKLEMAVFIHHLVLFKNMELADDDPFAFPVDFPENG 504
Db 431 -----YVLPFGNGPRNCIGRMFALMNMKALTKVLQNFSPQCKETQ----- 473
QY 505 LPIRVS 511
Db 474 IPLKISR 480

RESULT 9
PCT-US95-05758-25
Sequence 25, Application PC/TUS9505758
GENERAL INFORMATION:
APPLICANT: Cornell Research Foundation, Inc.
TITLE OF INVENTION: Cytochrome P4501pr Gene and Its
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05758
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1304
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rat
STRAIN: Unknown
DEVELOPMENTAL STAGE: Adult
POSITION IN GENOME:
CHROMOSOME/SEGMENT: Unknown
PCT-US95-05758-25

Query Match 9.8%; Score 263.5; DB 5; Length 504;
Best Local Similarity 24.1%; Pred. No. 8.1e-18;
Matches 132; Conservative 89; Mismatches 213; Indels 113; Gaps 25;

QY 9 LLLPLLLSLSLFLILLKRRNKRTR--FNLP--PKSGMPELGETIGYLPKPTATTL 64
Db 3 LLSALTEHTWLLAVLVLLVGLFGTRTGLRKKQGIPEPKPLPFGVGLNLY-----MGL 57
QY 65 GDFMOQHSSKTKYIRSLNF--GE--PTIYSADAGLNRFLQNGRLFEC-----STPRSIG- 117
Db 58 WKFDVECHKKKYKTIW--GLFDGOMPLFAITTEMTKNVL-----VKECFSVFTNRDGP 110
QY 118 -GILGKMSGLVGVGMHMDMSISLNFSLHARLFTILKDVERTHLFVLDSDMOONS----- 172
Db 111 VGIMCKAIVSAKDEEMKRYRALLSTFTS--GRKE-MPTIIEQGDILVKKLQDAELGK 168
QY 173 -----IFSA--ODEAKFTENLAKHIMS--MDGEEETQLK--EYVTEMKGVSAPL 221
Db 169 PVTMKKVGAYSMVDYTSFQVNDVSLNPKDPVEKTKILRRDFDPLFSVLVLPF 228
QY 222 NLPGTAHYKAL-----QSRATILKFTERKMEERKLDIREDEDEEVEVKEDAEKMS 274
Db 229 LTP--IYEMLNICMFPRDSIEFKKFVYR--MKETRL----- 261
QY 275 DHVRKORTDDLLGWL-----KHSNLSTEQIIDLILSLFAGHETSSVAIALAI 324
Db 262 DSVQKHRV--DFLOLMNAHNDKSKDESHETALSDMEITAOISIFIFAGYEPTSSITLSFVL 319
QY 325 FFLQACPRAVELEHREHLEIARAKKEGESELMNDYKMKDFTQCVINETRLGNAVREL 384
Db 320 HSLATHPTQKKLOE---EIDRALP--NKAPPTYDTVEMEMELDMVLEITRLPIGRL 374
QY 385 HRKALKDVRKGYDIPSGWKVLPVISAVALDNSRYDQNLFPNPMWMOOONNGASSSGSGS 444
Db 375 ERVCKKDYEVGFMFKGSVMIPISYALHRDPQHWPPEEPERPERFSKENGSDP---- 430
QY 445 FSTGNNNTMPFGGPRLCAGSELAKLEMAVFIHHLVLFKNMELADDDPFAFPVDFPENG 504
Db 431 -----YVLPFGNGPRNCIGRMFALMNMKALTKVLQNFSPQCKETQ----- 473
QY 505 LPIRVS 511
Db 474 IPLKISR 480

RESULT 10
US-08-313-075A-30
Sequence 30, Application US/08313075A
Patent No. 5639670
GENERAL INFORMATION:
APPLICANT: Holton, Timothy A.
APPLICANT: Cornish, Regina C.
APPLICANT: Tanaka, Yoshikazu
TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,075A
FILING DATE: 30-NOV-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PL 1538/92
FILING DATE: 27-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PL 6698/93

```

?      FILING DATE: 07-JAN-1993
?      PRIOR APPLICATION NUMBER:
?      APPLICATION NUMBER: AU PCT/A093/00127
?      FILING DATE: 25-MAR-1993
?      ATTORNEY/AGENT INFORMATION:
?      NAME: Digilio, Frank S.
?      REGISTRATION NUMBER: 31,346
?      REFERENCE/DOCKET NUMBER: 9433
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: (516) 742-4343
?      TELEFAX: (516) 742-4366
?      TELEX: 230 901 SANS UR
?      INFORMATION FOR SEQ ID NO: 30:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 476 amino acids
?      TYPE: amino acid
?      TOPOLOGY: linear
?      MOLECULE TYPE: protein
?      OS-08-313-075A-30

```

```

•Query Match Similarity 9.68; Score 257; DB 1; Length 476;
  Best Local Similarity 24.08; Pred. No 3.3e-17;
Matches 123; Conservative 87; Mismatches 227; Indels 76; Gaps 21;

```

```

0Y 12 LLLPILSLLELLPILKKRRKRRFLPRGKSGMPLSTGTIGYIKRYATATLTGDMQOH 71
Db 9 LLIIFSLISLTKM-----KSNCOIK-KLPGRPKVPLPGLSLHMMVGGPLPHNVLRLD---- 59
QY 72 VSKYKGIYRSLNFEPTIVSADAGLNRFILONEGRLECSY-PRSIGIIGKSMVLV- 129
Db 60 AKKQPIHNLÖLGISAVVVVSPEMARKVPLKTHDLAF--AYRPKLIGIEIVCYNSDIAF 117
QY 130 ---GDMHRDMRSIS-LNPLSHARLRTILLDNVEHTLFLVDSMQONS-----ITS 175
Db 118 SPYGDYMRÖMKICVLEYLSKKNRS--FNSIRDELIMIDPLRSSSGKPRVITERIES 175
QY 176 AÖDEAKKRTFNLMAKHIMSMDPGE-EETEOUKREYVFMKGVSAPINLTGTAY-HKALO 233
Db 176 -----FTSMICRSYFGKRIKEDKDECIRHVAKMTGLIGDFVADI-PSLRFHLVLIG 227
QY 234 SRATILKIEKKMERKLDIKEBOÖEEVEKYTEDAEAMSKSDHVRKQRTDOLLGWLKH 233
Db 228 MKGKIMD-VHKV-----DAIVEYVNNER-FTLTGTGTNGVGGEDLIDLVLRL 275
QY 294 SN-----LSTEOILDLILSLFAGHETSVALAIAIFLOACPKAVEELREHLEIAR 346
Db 276 KEEDGLOLPITNDNIKAIIFNDMFAGTSTSTINMAMVELMKNPVSFAKAOÖEVREYVK 335
QY 347 AKKELGEBELMWDYKKMDPFOCYINETLKL-GNVVAFILRKALKDVRKYGYDIPSGWKY 405
Db 336 GKFEFDE---DDIEELNYKLVILRETLRLHPLRLPLLRRECRRETEIINGYIPIPLTKV 390
QY 406 LPVTSVAHLNDSRYDOENLEFPMFMÖOONNGASSGSGSFTWNGN--YMFEGGAPLCA 463
Db 391 IVNWVALIGROPKYMDDAESFPERFEH-----NSLNFGNNEVELPFGSGRIICP 440
QY 464 GSELAKEMAVFIHLLVLKFWMEALBEDDÖPPAF 496
Db 441 GISFGLNAVYHPLAÖLLYHNEFMRPLRPTGVDPNDF 473

```

```

RESULT 11
US-08-201-118-7
: Sequence 7, Application US/08201118
: Patent No. 5786191
:
: GENERAL INFORMATION:
:
: APPLICANT: GOLDSTEIN, Joyce A.
: APPLICANT: ROMKES-SPARKS, Marjorie
: TITLE OF INVENTION: CLONING AND EXPRESSION OF COMPLEMENTARY
: TITLE OF INVENTION: DNAS FOR MULTIPLE MEMBERS OF THE HUMAN CYTOCHROME P450
: TITLE OF INVENTION: SUBFAMILY
: NUMBER OF SEQUENCES: 44
:

```

1 CORRESPONDENCE ADDRESS:
2 ADDRESSEE: Townsend and Townsend Kourlie and Crew
3 STREET: 379 Lytton Avenue
4 CITY: Palo Alto
5 STATE: California
6 COUNTRY: US
7 ZIP: 94301
8
9 COMPUTER READABLE FORM:
10 MEDIUM TYPE: Floppy disk
11 COMPUTER: IBM PC compatible
12 OPERATING SYSTEM: PC-DOS/MS-DOS
13 SOFTWARE: Patent In Release #1.0, Version #1.25
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/08/201,118
16 FILING DATE: 22-FEB-1994
17 CLASSIFICATION: 435
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: US 07/864,962
20 FILING DATE: 09-APR-1992
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Liebeschuetz, Joe
23 REGISTRATION NUMBER: 37,505
24 REFERENCE/DOCKET NUMBER: 15280-192-1
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: (415) 326-2400
27 TELEFAX: (415) 326-2422
28 INFORMATION FOR SEQ ID NO: 7:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH: 490 amino acids
31 TYPE: amino acid
32 STRANDEDNESS: single
33 TOPOLOGY: linear
34 MOLECULE TYPE: protein
35 US-08-201-118-7

Query Match	9.6%	Score 256.5	DB 1	Length 490
Best Local Similarity	20.8%	Pred. No. 3.9e-17		
Matches 110; Conservative	99	Mismatches 202	Indels 119	Gaps 20

```

0Y 11 PELLRLPILLSLLLELLKRNKRKTRNLPKRGKGMFLSETI-----GLKPYTATTLG 65
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 PFVVALVCLSLPMLFSLMRQSCRRRK--LRPGPRPLRTIGMMQIDVKDICKST----- 55
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 66 DEMOONHVK-YGKIYRNLFEPTITVSADAGLNFITLONEGRJRECSYPRSIGILLK-W 123
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 56 -----NEKVYGVGFYTFYFGMNP LVEFHGVAIVKAEALDNGEESHSGNSNPISORITKGL 110
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 124 SMLVLVDGMHDMRSISLNLSHARLRTILKD--VERHTL-----FVL 165
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 111 GIISNGKRMKEIRFESLTNLNFGMGKRSIEDYVOEALHVLBELKRTASPDPPFYL 170
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 166 DSMOONSFEKSODEAKKRTTFYMLAKHMSMDPGBEETLOLKKEVYFRMK-----GVSA 219
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 GCAPCNVLCISVYFC-KRDYK-----DQNLTLMKFNFENFRILNS 210
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 220 PL-----NLPGTAVHAKLQSHAPTLKTEKRMKEERKIDIKEEDOESEE---V 263
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 211 PWIOVCNPFPLIDOCPEPT-HNKKLVKVALYLRSYIRKEVYENQASLDVNNPRFMDCFLI 269
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 264 KTEDEAEKSKSDHNRKQKOTDDDLGWLKHNLSLSTEQIDLDLSLRFAGHETSVALALA 323
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 270 KMEQCKDKQKSE-----FNIENLVGTADLFVAGTETSTTYLRG 309
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 334 IFELDACPRAVEELREE--HLEIARAKKELESESLNMDVYKKADFTOCVINTLRLGAV 381
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 LLLLLKHEVLAKEVOEETIDHW-IGRHSRSC-----MQRSHMPYUADAVVNEIORYSDLV 362
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 382 -RFLRKRLKDYRKGYDIPRSGMKVULPYIYSAVHLNDSRYOQPNLFNFMWQOONNGSSS 440
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 363 PTCVPHAVTTPCKFRNVLIPRGTITMALLTSVLHDDDEFPRNPIDVDEGHFLDNK----- 416
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

0Y 441 GSGSPSTVNGNNVMEFGGGRPLCAGSELAKTEMAFVHNLVLKFMWEAEAD 490

```

Db 417 -GNFKK-SDYFMPFSAGKRICAGEGLARMELFLETTIIONFNLSKSVDD 463

RESULT 12
US-08-238-821B-7
Sequence 7, Application US/08238821B
Patent No. 5912120
GENERAL INFORMATION:
APPLICANT: GOLDSTEIN, Joyce A.
APPLICANT: ROMKES-SPARKS, Marjorie
TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN
TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S-
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/238,821B
FILING DATE: 06-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/201,118
FILING DATE: 22-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,962
FILING DATE: 09-APR-1992

ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 15280-192110US
TELEPHONE: (650) 326-2400
TELEFAX: (650) 326-2422
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 490 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-238-821B-7

Query Match 9.6%; Score 256.5; DB 2; Length 490;
Best Local Similarity 20.8%; Pred. No. 3.9e-17;
Matches 110; Conservative 99; Mismatches 202; Indels 119; Gaps 20;

Db 11 PLLLPSSLFLFLILKRRNRKTRNLPKSGMFLGFTI-----GLKPYTATTLG 65
Db 3 FVYLVLCISFLMLESFIMROSCRRK--LPPGPPPLPIINMLOIDVKICKSFT----- 55
QY 66 DFMQOHVSK-YGKIYRNSLGEPTIYSADAGLNRFLQNGRGLFECSYPSIGILGK-W 123
Db 56 -----NSKYVGPVFTYVPGMNPVIVHGEAYKEALIDNGEESGSGNSPIQRIITKGL 110
QY 124 SMALVYGDHMDRMSISLNLFLSHARLTITLLKD---VERHTL-----EVL 165
Db 111 GIISNCKRKKEIRRFSLTILNRNFGMKRSIEDRVOEAEHCLVEELKRTKASPCDPFFIL 170
QY 166 DSMQONSIFSAODEAKKFTFNLAKHIMSDPBEETEDJLKKEYVTFPMK-----GVVSA 219

Db 171 GCAPCNVCISVWQ-KRFDPK-----DQNFLLTKRFFENFRILNS 210
QY 220 PL-----NLPGTAVKALQSRATILKFERKMEERKLDIKEEQEE-----V 263
Db 211 PMIOVCNNPFLIDCFEFGT-HNKYLKNAVLTSTRYIKRKYAEHQASLDVNNPRDPMGCFLI 269
QY 264 KTEDEAEMSKSDHYRKQRTDDLLGVNLKHSNLSTEQIILILSLFAGHETSVAIALA 323
Db 270 KMEDEKDNQKSE-----FNENILVGTADVLEFVAGTETSTILRYG 309
QY 324 IFPLQAPKAVEELRER-LEIETAKKELGESESLMNDVKKKDFQCVINETLRGNV 381
Db 310 LLLLKHPEVTAKVQEEIDIV-IGRRSPC-----MQRSHMPTDVAVHEIQRSIDL 362
QY 382 -RFLHRKALDVRKYGYDIPSGWKVLPVISAHLDSRYDQPLFNFRMQOONCASS 440
Db 363 PTGVPAHVTTDTKFRNRLIFKGTIMALLSVLHDDKEFPNPIFDGCHLDKN----- 416
QY 441 GSGSFSTWGNMYPFGGPRLCAGSELAKLEMAVFIIHLVLFKNWELAE 490
Db 417 -GNFKK-SDYFMPFSAGKRICAGEGLARMELFLETTIIONFNLSKSVDD 463

RESULT 13
PCT-US95-05744-7
Sequence 7, Application PC/TUS9505744
GENERAL INFORMATION:
APPLICANT: GOLDSTEIN, Joyce A.
APPLICANT: ROMKES-SPARKS, Marjorie
TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN
TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05744
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/238,821
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/201,118
FILING DATE: 22-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,962
FILING DATE: 09-APR-1992

ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 15280-192-1-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 490 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
PCT-US95-05744-7

Query Match 9.6%, Score 256.5, DB 5, Length 490;
Best Local Similarity 20.8%, Pred. No. 3.9e-17;
Matches 110; Conservative 99; Mismatches 202; Indels 119; Gaps 20;

QY 11 PLLLPILLSLLPILILKRRRRKTRFNLPPKSGMPFLGEGT-----GYLKPYATTILG 65
DB 3 PFVVLVLCLSLFELSLRQSCRRK--LPPGPPLPIGNNLQIDVCKISFT----- 55
QY 66 DEMOQHVK--GKTIYSMLFEGPTIVSADAGLNRFILONEGRLEFECSPRSIGILGK-W 123
DB 56 -----NFKVIGPVTVYFGAMPYIVFEGYEAVKCALINDGEEFGSGRNSPIQRITGL 110
QY 124 SMLVLVGMHMRDRSISLNFSLHARLRITLKD---VERHTL-----FVL 165
DB 111 GLISNGKRMKEIRRFSLTNLNFQMKRSIEDRYOEENHCLVEELRKTAKSPCDPTFL 170
QY 166 DSMOONSIFSAODEAKKFTFNLMAKHIMSMDGEEETOLKKEVYTFMK-----GVYSA 219
DB 171 GCAPCNVCSVVEQ--KRDYK-----DONFLTLMKRFENEFRLNS 210
QY 220 PL-----NLPGTAVHKALOSRAITLFIERKMEERKLDIKEEDQEE-----V 263
DB 211 PWIOVCNNPRLIDCFPGT--HNKVLKNAVALTRSYIREKKEHQASLDVANNPDMDFLI 269
QY 264 KTEDEAKSKSDHVRKQRTDDDLGWLKHSNLSTEQILDLILSLFAGHETSSVAIALA 323
DB 270 KMEQERKDKSE-----FNIEHLVGTAVDLFVAGTETSTTLRYG 309
QY 324 IFFLOACRAVELREE--HLEIARAKKELGESELMNDYKMDTQCINTELIGNV 381
DB 310 LLLLKHPEVYAKVQEEIDH--IGHRSPC-----MODRSMPTDAVHIEIQRSIDL 362
QY 382 -RFLRKALDKVRYKGYDIPSGMKVLPVISAHLNDRYDQPNLFNPMRMOQONNGCASS 440
DB 363 PIGVHAHTYTDKFRNYLLPKGTITMALLTSVLHDKKEPNINIDPGHFLDK----- 416
QY 441 GSGSTYNNMYPGGGPRLCAGSELAKLEMAVFIHHLVLFKNMELAD 490
DB 417 -GNFKK--SDYFMPFSAGKRICAGELARMELFLFTLILQNFNLKSVSD 463

RESULT 14
US-08-532-065B-2
Sequence 2, Application US/08532065B
Patent No. 5753507

GENERAL INFORMATION:
APPLICANT: Ohta, Daisaku
APPLICANT: Mizutani, Masaharu
TITLE OF INVENTION: Plant Geraniol/Nerol 10-Hydroxylase and
TITLE OF INVENTION: DNA Coding Therefor
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5753507artis Corporation
STREET: 59 Route 10
CITY: East Hanover
STATE: NJ
COUNTRY: USA
ZIP: 07936
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,065B
FILING DATE: 22-SEP-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 495 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-532-065B-2

Query Match 9.1%, Score 245, DB 1, Length 495;
Best Local Similarity 22.9%, Pred. No. 5.6e-16;
Matches 120; Conservative 80; Mismatches 218; Indels 106; Gaps 19;

QY 12 LLLLSLSL--LLPILILKRRRRKTRFNLPPKSGMPFLGEGTGYLKPYATTILGDFMQQ 70
DB 10 LLLPFIILSCFLFTTTSGRISRGATALPPGPPLPIG--NIHLVGNHPRHSFALSK- 67
QY 71 HSKYIGKTIYSMLFEGPTIVSADAGLNRFILONEGRLEFECSP-----RSIG----- 117
DB 68 --TYGPWMSLKGSLNIVVIASPEAREVLRTHQIILSARSPTVAVRSINHQDASLWVL 124
QY 118 -GILKWSMLVGMHMRDRSISL--NFLSHARL---RTILKQVERHTLPVLDQMONS 172
DB 125 PSSSARWRL-----RRLSVTLQSLPQRIENTATLRNKKVKELVSFSESSDRE 174
QY 173 IFSADDEAKKFTFNLMAKHIMSMDGEEETOLKKEVYTFMKGVASAPLNLPGT----- 226
DB 175 SVDISRVAFTITLILNISLFSVDLSYNA-----KASINGVDYIVSWDADAGTDAANY 230
QY 227 ---AYHKAQSRAITLK-----FIERKMEERKLDIKEEDQEEVKTDEAEM 271
DB 231 PFPLRFILQGVNKTFCYCTERLVVFRGFIIDAKIAEKSSQNNPKD-----V 277
QY 272 SKSDVVRKQRTDDDLGWLKHSNLSTEQILDLILSLFAGHETSSVAIALAIFFLQACP 331
DB 278 SKNDYV-----DNLDDKGESELSISDIEHLDMFTAGTIDSSSTLEMPTELLKNP 331
QY 332 KAVEELREHEHLEIARAKK-----LGESEL--NMDDYKMDFTQCVINETLRIGNVVRL-H 385
DB 332 KT-----MAKAQAEIDCVIGONGIVEESDISKLPYLGAVVKEPRLRTPVLLIP 381
QY 386 KRALNDVRKGYDIPSGMKVLPVISAHLNDRYDQPNLFNPMRMOQONNGSSGSGSF 445
DB 382 RKAESDAEILGPMVLKDVQVILVNWVAIGRDPVWNPQSFEEERFLGKD-----M 431
QY 446 STWGNVY--MPGGGPRLCAGSELAKLEMAVFIHHLVLFKNMEL 487
DB 432 DVRGRDIELTPFGAGRRICPGMPLAMKTVSLMALSLIYSFDKRL 475

RESULT 15
US-08-948-564-16
Sequence 16, Application US/08948564
Patent No. 6121512

GENERAL INFORMATION:
APPLICANT: Siminszky, Balazs
APPLICANT: Dewey, Ralph E.
APPLICANT: Corbin, Frederick T.
TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-854-1400
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-948-564-16

Query Match 9.18; Score 244.5; DB 3; Length 576;
Best Local Similarity 22.7%; Pred. No. 7.9e-16;
Matches 111; Conservative 91; Mismatches 181; Indels 107; Gaps 22;

QY 63 TLGDFMOQHYSKYGIYRSLNFGPTIVSADAGLNRFILQ-----NEGRLEFGESYPRSI 116
DB 97 SLIDWFELEH-----GAVYKLFPGKAFYVSDPIYARHILKRENAFSYDKGLADILEP--- 149
QY 117 GILGK-----WSM--LVLVGDMHRMRSISLNLSHARLTILL--KDYERHTLF 163
DB 150 --IMGKGLIPADLTDTWKORRRVIAPAFHNSYLEAMVKITTCSERTILKFNKLLGEGYD 207
QY 164 VLDSMOONSIFSAODEAKKFTFNMAKHISMOPGE--EETEOUKREYVFPMKGVASAPL 221
DB 208 GPDSIE---LDLEAEFSSSLALDIIIGLVFNYPDGSVTKESPVIKAVYGTLFEAHRSTF 263
QY 222 NLPGTAYHKAQSRAFLIKFTERKMEERKLDIK-----BEDQEEEEVKTEDE 268
DB 264 YIP---YWKIPLARWIVPR--QKRFGD--DLKYINTCLDGLIRNAKESRQETDVE---- 311
QY 269 AEMSKSDHVRKQRTDDLLGMV--LKHNSLSTEQIIDLILSLFAGHETSSVAIALAIF 326
DB 312 -KIQQRDYLNLIK--DASLLRFLVDMRGADYDDRLRDLMTMLAGHETTAAYLTWAVFL 368
QY 327 LQACPKAVEELREHLEIARAKKELGESELMNDYKKMDFTQCIYINETLRGNVYRFLHR 386
DB 369 LAONPSKMKRAQ-----AEVDLVLTGTRPFESLKELOYIRLIVEALRLYPOPLLIR 422
QY 387 KALK-DV-----RYKGYDIPSGMKVLPYISAVHLDSRY--DQPNLFNPMWMOQONG 436
DB 423 RSLKSDVLPGGHKEGKGYALPAGTDVF--ISVYNLHRSPYFMDRPPDFEERFLVQKN 480
QY 437 ASSSGSGSESTWGN-----NYMPFGGPRICAGSELAKLEMAVFIH 477
DB 481 EEIEG-----WAGLDPSRSPGALYPNEVISDFAFLPFGGGPRKCVGDQFALMESTVALT 534
QY 478 HLYLKFNNMEL 487
DB 535 MLQONFDVEL 544

Search completed: September 28, 2001, 18:44:39
Job time: 69 sec

PS Claim 50: Fig 11; 113pp; English.
 XX
 CC The present sequence represents a DWf4 polypeptide. The polypeptide is a
 CC cytochrome P450 enzyme that mediates multiple steps in synthesis of
 CC brassinosteroids. Specifically, it mediates multiple
 CC 22alpha-hydroxylation steps in brassinosteroid biosynthesis. The DWf4
 CC polynucleotide is used for altering the phenotype of a plant. DWf4
 CC plants display a dramatic reduction in the length of different organs,
 CC and this size reduction is attributable to a defect in cell elongation.
 CC The DWf4 polynucleotides and polypeptides can be used in diagnostic
 CC assays and to generate antibodies, which can be used to produce
 CC immunogenic compositions.
 XX
 SQ Sequence 513 AA:

Query Match 100.0%; Score 2681; DB 21; Length 513;
 Best Local Similarity 100.0%; Pred. No. 6.6e-229;
 Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFEYEHHTLLPPLLPLSLLLFLILLKRRNRKTRFNLPPGKSGWPLGTTGYLKPYT 60
 DB 1 mfeYehhtllp1lllpslllfl1lllkrnrktrfnlppgksgwpllgeltgylkpyt 60
 QY 61 ATTLDGMQOHVSKYKGYKIVSNLGEPTIVSADAGLNRFILQNGRLEFCSYPRSIGIL 120
 DB 61 attldgmqohvskykgykivsnlgeptivsadeqlnrflilqngrlfecsyprrsigil 120
 QY 121 GKMSMLVLCMDHMRDMSISLNFSLHARLRTILKDYERHTLFVLDSMOQNSIFSAQDEA 180
 DB 121 gkmsmlvldcmdhmrmsislnfslharrtllkdyerhtlflvldsmoqnsifsaqdea 180
 QY 121 gkmsmlvldcmdhmrmsislnfsharrtllkdyerhtlflvldsgqnsifsaqdea 180
 DB 181 KKTFTNMAKHMSMDGEEETEOLKKEYVTFMKGVVSAPLNLPGTAYHKALOSRAITLK 240
 DB 181 kktftnmaKhmsmdgeeteqlkkeyvtfmkgyvsaplnlpgtayhkaIgsratilk 240
 QY 241 FIERKMEERLDTKEEDQEEEVKTEDEAEKMSKSDHVRKORTDDLLGWVLKHSNLTSEQ 300
 DB 241 fIerKMEerldtkeedqeeevktedeaemsksdhvrkgrtdddlIgwvlkhsnltseq 300
 QY 301 ILDLILSLFAGHETSSVAIALAIFLQACPKAVEELREHELETARAKKEGSEELNMWD 360
 DB 301 lldlilslfaghetssvaialaifflqacpkaveelreehleIarakkelgeeselnwd 360
 QY 361 YKKMDFTQCVINETLRIGNVVRFLLHRAKLDVRYKGYDIFSGWKVLPVISAVHLNDRYD 420
 DB 361 ykkmdftqcvinecltrignvvrflhrkaldvrykgydipsgwkvlpvisavhlnDnsryd 420
 QY 421 OPNLFNWRMOQONNGASSSGSFSFTWGNVYMPFGGPRLCAGSEELAKLEMAVFTHHLY 480
 DB 421 opnlfnwrmoqonngasssgsfstwnvymfpggprlcagseelaklemavfthhly 480
 QY 481 LKFNWELAEDDQFPAPFPVDFPNGLPIRVSRIL 513
 DB 481 lknfwelaeddqpfapfpvdfpnglpirvsrill 513

RESULT 2

AAW27153 standard; Protein; 472 AA.

AAW27153:

14-APR-1998 (first entry)

Arabidopsis thaliana cytochrome P450-type hydroxylase.

KM Cytochrome P450-type hydroxylase; identification; brassinosteroid;
 KM brassinosteroid inhibitor; modified plant; recombinant production;
 KW testosterone.
 XX
 OS Arabidopsis thaliana.
 XX

PN W09735986-A1.
 XX
 PD 02-OCT-1997.
 XX
 PF 27-MAR-1997; 97MO-EP01586.
 XX
 PR 27-MAR-1996; 96US-0622166.
 XX
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 XX
 PI Altmann T, Koncz C, Mathur J, Szekeres MA;
 XX
 DR WPI; 1997-489649/45.
 DR N-PDB; AAT85306, AAT85307.
 XX
 PT New isolated plant cytochrome P450-type hydroxylase gene - used to
 PT identify substances acting as brassino-steroid(s) or brassinosteroid
 PT inhibitors for the production of modified plants
 PS
 PS Claim 1: Pages 44-46; 77pp; English.

CC The present sequence is Arabidopsis thaliana cytochrome
 CC P450-type hydroxylase. The hydroxylase can be used to identify
 CC brassinosteroids or brassinosteroid inhibitors, useful to produce
 CC plants with modified physiological and/or phenotypic
 CC characteristics. The modified plants may show, e.g. stimulation of
 CC growth, increased cell elongation, increased wood production,
 CC accelerated seed germination at low temperatures, an increase in
 CC dry weight, repressed anthocyanin production during growth in light
 CC and/or inhibited de-etiolation which is induced, e.g. by cytokinin,
 CC in the dark or an increase in stress tolerance. The hydroxylase or
 CC its coding sequence can also be used for the recombinant production
 CC of compounds, e.g. testosterone.
 CC
 SQ Sequence 472 AA:

Query Match 36.5%; Score 978.5; DB 18; Length 472;
 Best Local Similarity 41.9%; Pred. No. 3e-78;
 Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;

QY 12 LLLPLSLLLFLILLKRRNRKTRFNLPPGKSGWPLGTTGYLKRYTATLGDPMQOH 71
 DB 7 llllpslllfl1lllkrnrktrfnlppgksgwpllgeltgylkrytatlgdpmoqh 71
 QY 72 VSKYKGYKIVSNLGEPTIVSADAGLNRFILQNGRLEFCSYPRSIGILGKMSMLVLCMD 131
 DB 72 vskykgykivsnlgeptivsadaglnrfilqngrlfecsyprrsigilgkmsmlvldc 131
 QY 132 MHRDMSISLNFSLHARLRTILKDYERHTLFVLDSMOQNSIFSAQDEAKKFTNMAKH 191
 DB 132 mhrdmsislnfslharrtllkdyerhtlflvldsmoqnsifsaqdeakkftnmaKh 191
 QY 192 IMSMDGEEETEOLKKEYVTFMKGVVSAPLNLPGTAYHKALOSRAITLKFIERKMEER 251
 DB 192 imsmdgeeteolKkeyvtfmkgyvsaplnlpgtayhkaIgsratilkfIerKMEer 251
 QY 252 DIRKEDQEEEVKTEDEAEKMSKSDHVRKORTDDLLGWVLKHSNLTSEQILDLILSLFA 311
 DB 252 dirkedqeeevktedeaemsksdhvrkortddllGwvlkhsnltseqilDlilslfa 311
 QY 312 GHETSSVAIALAIFLQACPKAVEELREHELETARAKKEGSEELNMWDYKKMDFTQCVI 371
 DB 312 ghetsSvAialaifflqacpkaveelreehleIarakkelgeeselnwdYkkmdftqCvi 371
 QY 372 NETLRIGNVVRFLLHRAKLDVRYKGYDIPSGWKVLPVISAVHLNDRYDOPNLFNWRMO 431
 DB 372 netlriGnvvrflhrkaldvrykgydipsgwkvlpvisavhlnDnsrydOpnlfnwrmo 431
 QY 432 QONNGASSSGSFSFTWGNVYMPFGGPRLCAGSEELAKLEMAVFTHHLYLKFNMWELAE 490
 DB 432 qonngasssgsfstwnvymfpggprlcagseelaklemavfthhlylKfnmwelaed 490
 QY 490 -----snvltgprsnvfpfsggprlcpgyelaaralavflnrlvtgfsawypaeq 446
 DB 490 -----snvltgprsnvfpfsggprlcpgyelaaralavflnrlvtgfsawypaeq 446

OY 491 DQFAFPVDFPNCGLPIRYSR 511
| : | | | |
DB 447 dklvffplrtqtkypifvkr 467

RESULT 3
AAG44571
ID AAG44571 standard; Protein: 472 AA.
XX
AC AAG44571;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55847.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PE EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

```
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150066.
PR 26-AUG-1999; 990S-0150084.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0156559.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161922.
PR 28-OCT-1999; 990S-0161933.
PR 29-OCT-1999; 990S-0162142.

Query Match 36.5%; Score 978.5; DB 21; length 472;
Best Local Similarity 41.9%; Pred. No. 3e-78;
Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;
```

```
OY 12 LLLPSSLILFLILKRRNRKTRFNLPPKSGMPFLGFTIGYIKRYTATTTIGDEFQOH 71
DB 7 llllssaaafll---lllrrtyrrmgllppssjgllpilletfgllgasykxtenpeptlde 63
OY 72 VSKYGGKIYRNLGCEPTIVSADAGLNRFLIQNNGRLFECSYPSRISGILGKWSMLVLVGD 131
DB 64 varygsfwmhllgpeplfsadpetnrfvlnqnegklfecsyasiscnllgkhslllmkgs 123
```

```
OY 132 MHRMRSISLNFSLHARLRTLLKDYERHTLFVLDSSQONSISFSAODEAKKFTFNLMK 191
DB 124 lhrkmhslmsfanssllkohlmdlgrlvrfnldssvsvll--meaakkltfelvkq 181
OY 192 IMSMDPEEETEOILKRYVYFMKGVSAAPLNLPGTAYHKLQSRATLTKFERKMEERKL 251
DB 182 lmsfdpg-ewseslrkeyllviegffsrlprlfstlyrkaigar-----fkvaecalt 232
OY 252 DIKEEDOESEEVKTEDEAKMSKSDHVAKKORTDDLLGWLKHSLSPEOILLDLSLFA 311
DB 233 vvmkkrreeegae-----fkkmllaalldaaddgfsdeeltvdlvaliva 278
OY 312 GHETSSVALAIATFLOACPKAVEELREEHLEIARAKKEGESELMDDYRKMDFTQCVI 371
DB 279 gyetstlmvlavkflletpqlalaqlkeehketrakmsd--sytlewsdykmpftgcgv 336
OY 372 NETRLGCVYRFLHRLKALKVRYKGYDIPSGWKVLPYISAVHLDNSRYDDPNLFPWRMQ 431
DB 337 netlrvanllgvyfiramtvelkykllprgwkvyfssfravhldpnfkdarlnprwtq 396
OY 432 QONNGASSGSGSFSTWGN--YMPFGGPRLCAGSELAKEAMVFHHLVLKFMWELAED 490
DB 397 -----snsvltgpsvnltpcfpg99pfcipyelatarvalsvflnlvtgfswpaeq 446
OY 491 DOPFAFPFVDFPNGLPIRVSR 511
DB 447 dkivffptltqkrpflvkr 467

RESULT 4
AAG45022
ID AAG45022 standard; Protein; 472 AA.
XX
AC AAG45022:
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56469.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
FE 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 23-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
```

PR 07-MAY-1999;	99US-0132863.
PR 11-MAY-1999;	99US-0134256.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134219.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 10-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 17-JUN-1999;	99US-0139453.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 02-JUL-1999;	99US-0142154.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 26-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 16-AUG-1999;	99US-0148684.
PR 17-AUG-1999;	99US-0149368.
PR 18-AUG-1999;	99US-0149175.
PR 20-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159263.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 18-OCT-1999;	99US-0159638.
PR 21-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

36.5%; Score 978.5; DB 21; Length 472;

Best Local Similarity 41.9%; Pred. No. 3e-78;

Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;

QY 12 LLLPPLSLFLILKRRNRKTRNLPBGKSGWPLGEGTIGLYKPYATTLGDFMQOH 71
DB 7 llllsiaagfl---lllrtyrmglppgsigllpigtftqllgkyktenpepfider 63
QY 72 VSKYGRIVSNLGEPEPTIVSADAGLNRFILQNGRLFECSYPRSIGIGIGKSMGLVGD 131
DB 64 varygsvfmcchlgepctllsaopetrfrvlgneqkllfcsypsalcnlllgkshlllmkgs 123
QY 132 MHRDMRSISLNFSLHARLRTLLKDYERHTLFVLDQMQNSIFSADDEAKKFTFNMAKH 191
DB 124 lhrmrlmslmsfnsllkdhlmldldrlvrldswsvll--meakkltfcltvkg 181
QY 192 INSMDDGEETEOLEKKEVYTFMKGVVSAPLNDPTAYHKAQSRATILKFERKMEERKL 251
DB 182 lmsfdpg-ewseslrkeyllvlegfslpplfstctyrkaigar-----rkvaealt 232
QY 252 DIKEOEDEEVVTEDEAEKMSKSDHVRKORTDDLLGWVLKHSNLSDEOILLLSLFLPA 311
DB 233 vvmkrrreeeeagae-----rkkmlaallaaadgfdedeivdlvaliya 278
QY 312 GHETSSVAIALAIFLQACRAVEELREHLEIARAKKELGESELMWDDYKKMFTQCVI 371
DB 279 gyetctatmlavkflteplrlaqllkeehkrlranksd--syslswdsdkmpfcqcv 336
QY 372 NEFLRGVNVRFHRRKALMDVRKGYDIPSGWKVLPVISAIVHNDNSRYDDPNLFNWRMQ 431
DB 337 neclrvanllgyvframctveikykllpkgwkvsvsfavhldphfkdaftmptwrg 396
QY 432 QQNNGASSSGSFSWGN--YMPFGGPRILCAGSEIATLEMAVFTHHLVLEKFNWIAED 490
DB 397 -----snsvtgpnavvfcpfggprlcpgyelarvalsvflhrlytgsfwpaeg 446
QY 491 DQPFAPFVDFPNGLPYRSR 511
DB 447 dklivftrtrtkryplfvkr 467

RESULT 5

AAG45021 standard; Protein; 491 AA.

AAG45021;

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 56468.

Protein identification; signal transduction pathway; metabolic pathway;

hybridisation assay; genetic mapping; gene expression control; promoter;

Arabidopsis thaliana.

EPI033405-A2.

XX

PD 06-SEP-2000.
XX 25-FEB-2000; 200DEP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139434.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.

```

PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149422.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

```

```

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 36.5%; Score 978.5; DB 21; Length 491;

Best Local Similarity 41.9%; Pred. No. 3.2e-78;

Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;

```

QY 12 LLLPSLSLLFLILKRNKTRFNLPGRSGWPFAGETGYLKPATTLGDPMOOH 71
DB 26 llllsiaagfl---lllrrtyrmglppsglpllgeltqlgavytentepefider 82
QY 72 VSKYKTYRNSLFGEPPIVSADAGLNRFLLONEGRLEFECSSYRSIGLIGKSMVLVGD 131
DB 83 varysvfmthlfgfeplfsadpelnrfvlgneqklfecspaslonllghkslllmkgs 142
QY 132 MRRDMRSISLNLSHARLRTILKQVERHTLVLDSSMOONSIFSAODEKAKKTFPNIMAKH 191
DB 143 lkhmhsilmsfianssilkdhmlldtrlvrlndlswssrvll--meeakkltfelvqk 200
QY 192 IMSMDPGEETPOLKKEYTTFPMKGVVSAPLNLPGYAVHKLQSRATILFKIRKMERLU 251
DB 201 lmsfdpg-ewseslirkeyllviegffslpblfstlytraiaqr-----rkvaeatl 251
QY 252 DIKEDQEEEEVKTDEDEAKMSKSDHVKRKORTDDDLGLWYLNKHSNUSTEOILDIISLFA 311
DB 252 vvmvkmrreeeeegae-----rkkdmiaallaadddgfsdeeiivdflvaliva 297
QY 312 GHETSVAIALAIFLQACPKAVEELREHNLIAARKKRLGSELNMDYKKMDFTQCVI 371
DB 298 gyeltslmltlavkltfletplalaqlkeehktrfmskd--syslewsdyksmpftqcvv 355
QY 372 NETLRGNVVRFLHRKALKDVRKYKGDIPIGSKVKVLPIVIAVHLDSRSRYDOPMLFNPWRMO 431
DB 356 netlrvanligvfrtramdvleikgklykpkvkvfsfvaivldphnfdatftfprwq 415
QY 432 QONNGASSSGSGSFSTWGN--YMPREGGPRLCAGSELAKLEMAVFIHHLVLFKNMELAD 490
DB 432 QONNGASSSGSGSFSTWGN--YMPREGGPRLCAGSELAKLEMAVFIHHLVLFKNMELAD 490

```

Db 416 -----snsvltgspnsvfcpfgggprlcpgyelarvalsvflhrvtgfswpaeq 465
Qy 491 DQPFAPFVDFPNGLPIRVR 511
|: | | |
Db 466 dkivfptlttqtkryplfvkr 486

RESULT 6
*AAG44570
ID AAG44570 standard; Protein; 492 AA.
XX
AC AAG44570;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55846.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0127462.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129645.
PR 19-APR-1999; 99US-0130047.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.


```
PR 13-AUG-1999; 99US-0148565-
PR 13-AUG-1999; 99US-0148684-
PR 16-AUG-1999; 99US-0149368-
PR 17-AUG-1999; 99US-0149175-
PR 18-AUG-1999; 99US-0149426-
PR 20-AUG-1999; 99US-0149722-
PR 20-AUG-1999; 99US-0149723-
PR 20-AUG-1999; 99US-0149929-
PR 23-AUG-1999; 99US-0149902-
PR 23-AUG-1999; 99US-0149930-
PR 25-AUG-1999; 99US-0150566-
PR 26-AUG-1999; 99US-0150884-
PR 27-AUG-1999; 99US-0151065-
PR 27-AUG-1999; 99US-0151080-
PR 30-AUG-1999; 99US-0151303-
PR 31-AUG-1999; 99US-0151438-
PR 01-SEP-1999; 99US-0151930-
PR 07-SEP-1999; 99US-0152363-
PR 10-SEP-1999; 99US-0153070-
PR 13-SEP-1999; 99US-0153758-
PR 15-SEP-1999; 99US-0154018-
PR 16-SEP-1999; 99US-0154039-
PR 20-SEP-1999; 99US-0154779-
PR 22-SEP-1999; 99US-0155139-
PR 23-SEP-1999; 99US-0155486-
PR 24-SEP-1999; 99US-0155659-
PR 28-SEP-1999; 99US-0156458-
PR 29-SEP-1999; 99US-0156596-
PR 04-OCT-1999; 99US-0157117-
PR 05-OCT-1999; 99US-0157753-
PR 06-OCT-1999; 99US-0157865-
PR 07-OCT-1999; 99US-0158029-
PR 08-OCT-1999; 99US-0158232-
PR 12-OCT-1999; 99US-0158369-
PR 13-OCT-1999; 99US-0159293-
PR 13-OCT-1999; 99US-0159294-
PR 14-OCT-1999; 99US-0159329-
PR 14-OCT-1999; 99US-0159330-
PR 14-OCT-1999; 99US-0159331-
PR 14-OCT-1999; 99US-0159637-
PR 14-OCT-1999; 99US-0159638-
PR 18-OCT-1999; 99US-0159584-
PR 21-OCT-1999; 99US-0160741-
PR 21-OCT-1999; 99US-0160767-
PR 21-OCT-1999; 99US-0160768-
PR 21-OCT-1999; 99US-0160770-
PR 21-OCT-1999; 99US-0160814-
PR 21-OCT-1999; 99US-0160815-
PR 22-OCT-1999; 99US-0160980-
PR 22-OCT-1999; 99US-0160981-
PR 22-OCT-1999; 99US-0160989-
PR 25-OCT-1999; 99US-0161404-
PR 25-OCT-1999; 99US-0161405-
PR 25-OCT-1999; 99US-0161406-
PR 26-OCT-1999; 99US-0161359-
PR 26-OCT-1999; 99US-0161360-
PR 26-OCT-1999; 99US-0161361-
PR 28-OCT-1999; 99US-0161920-
PR 28-OCT-1999; 99US-0161921-
PR 28-OCT-1999; 99US-0161993-
PR 29-OCT-1999; 99US-0162142-
Query Match 36.5%; Score 978.5; DB 21; Length 492;
Best Local Similarity 41.9%; Pred. No. 3.2e-78;
Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;
```

```
OY 12 LLLLSLILFLILIKKRNKTRNLPKSGMPLGELTIGYLKPYATITLGDPMOQH 71
DB 27 LLLLSiaagfL---Llrrttrrmgplppllgeltfqllyayktenpepfider 83
OY 72 VSKYKIRYSNLFGEPTIYSADAGINRFTLQNEGRLEFCSYPRSIGIIGKWSMLVQCD 131
```

```
DB 84 varysvfmthlfgpeptlfsadpecnrfvlgneqklfcesypasclnllgkshlllmgs 143
OY 132 MHRDMRSISLNFSLHARLRTILKDVERRHTLFVLDSMOONSIFSADOKAKKFTFNLMAKH 191
DB 144 lkhmrhslmsfanssllkhdldidrtvrfnlswserrvll--meaakktfelvq 201
OY 192 IMSMDGEEETEQLKKEVYTFEMKGVVASAPLNPGRAYHKAOSRATILKFERKMEERKL 251
DB 202 lmsfdpg-ewseslkeylivlegfslpdlpflstlyrkaiaqar-----rkvaalt 252
OY 252 DIKEEDOEEREVEKTEDEEMSKSDHVRKORTDDDLGLWKLKSNLSTEQIIDLILSLFLFA 311
DB 253 vvmkrrreeeegae-----rkkdmiaallaaddgdsdeeiavfvaillva 298
OY 312 GHETSSVALAIAlFPLQACPKAVEELREHLEIARAKKELGSEINMDYKMDFTQCVI 371
DB 299 gyetstlmclavkfllteplalaqlkeehkframskd--syslwsdykmpftgcgv 356
OY 372 NETLRGNVVRFLHRRKADVRYKGVDPISGKVLPIYSAVHLDSRXPQLFNPWRMO 431
DB 357 netlrvanllgvvfrtramdvaikgklypkykvtssfravhlldpnhkdartfnpwtq 416
OY 432 QQNGASSSGSCSFTWGNM-YMPFGGCPRLCAGSELAKEIAVFIHLVLKFNMLAED 490
DB 417 -----snsvtgpnsnvftpgfgygprlpgyelarvalsvflhrlvtgfsvpaeg 466
OY 491 DQFAPPEYDFPNGLPIRVSR 511
DB 467 dkLvfftrtqtgkryplfvkr 487
```

```
RESULT 7
AAG44572 standard; Protein; 444 AA.
ID AAG44572;
AC AAG44572;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55848.
DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
```

PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139464.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR- 01-JUL-1999; 9905-0141842.
PR- 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144086.
PR 16-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.

PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154039.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0155139.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 12-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.

RESULT 9
AA646490
ID AAC646490 standard; Protein: 461 AA.
XX
AC AAC646490;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58494.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144088.
PR 16-JUL-1999; 99US-0144088.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

```

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 27.1%; Score 726; DB 21; Length 461;

Best Local Similarity 32.3%; Pred. No. 6.6e-56; Mismatches 181; Indels 76; Gaps 10;

```

Matches 166; Conservative 91; Mismatches 181; Indels 76; Gaps 10;
QY 12 LLLPSSLISLLFLILKLRN--RKTRFNLPPCKSGMPFLGERTIGYKPTATITLGDFMQ 69
DB 7 LLLVSCSALL-----LWNGMYTKNGLPQTMGWPILGETTEFLKGP-----Nfmr 55
QY 70 QHVSQYKIRSNLFGERTIVSADAGLRFILONEGRFLFCSTYPRSTIGILGKMSMLVLV 129
DB 56 nqrlrygsffkshllgpcplismdsevrnyllkneskqlvpygysmdlllycmaavh 115
QY 130 GDMHRDMSISLFLSHARLRTLLKDVERRHTEFLVDSMOONSIFSAODEAKKETF---- 185
DB 116 gshrimrgslslisestumrthlpkvdhfmsyldqwelevidldqdktkmaflssl 175

```

```

QY 186 -----NLAKHIMSDPEEETEQLKKEYVFMKGVSAPLNLPSTAYHKAQSRATLK 240
DB 176 tqiagnlrkxpfv-----eefktafklyvgtslsvpidlpstnyrcslqgrmldr 225
QY 241 FIEKMEERKLDIEEDDEEEVETEDBAEMSKSDHYRKQRTDDLLGWLKHSN----L 296
DB 226 lreimqerr-----dsge-----tfdcmlylmkkegnrypl 258
QY 297 STEQILDLISLFPAGHETSSVALALAFIFQACPKAVEELREERHLEIARAKKELGSEEL 356
DB 259 tdeairqdvtillysyetvsttsmmalkyhdhpkalqelraehlafrerkrq--depl 316
QY 357 NMDYKKMDFTQCVINETLRLGNVVRFLHRRKALDVRKKGIDIPSGMKVLPVISAVHLDN 416
DB 317 glevdksmkfttravlyetstriatlingvlfktrtdleingyllpkwrylyytrelnyda 376
QY 417 SRVQPLNFNPMWQOONNGNASSGSGSFSWGNVNPFGGPRCLCAGSELAKEMAVFI 476
DB 377 nljedpdlfnpwrwmkkslesq-----nscfvfgygtrtlcpqkelgivelssfi 425
QY 477 HHLVKFNMELAEDDQPFAPFPVDFPGLPIRVS 510
DB 426 hylvtryweeiqgdelmvfrvtfapkgfhlris 459

RESULT 10
AAG46489
ID AAG46489 standard; Protein: 462 AA.
XX
AC AAG46489;
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58493.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 28-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.

```

PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0136021.
PR	25-MAY-1999;	99US-0136029.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	24-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	23-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	26-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
<hr/>		
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153738.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.

DB 430 hyfctiryweeioggdelmvpfrvafpkyghlrlis 463
RESULT 12
AAG11836
ID AAG11836 standard; Protein; 461 AA.
XX AAG11836;
-XX
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 10708.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.

```
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26.6%; Score 712; DB 21; Length 461;
Best Local Similarity 32.5%; Pred. No. 1,1e-54;
Matches 164; Conservative 93; Mismatches 196; Indels 52; Gaps 8;
```

```
DB 116 gpsnrlmgsllslstpsmmchlllpklddmrnylgwddetvdiqektkmf-1ss 174
QY 190 KHMSMPGEETEQLKKEYVTFMKGVSADPLNPGTAHYKALOSRATILKEIRKMEER 249
DB 175 llqiaetlkpeveeyreffecklvgtclsvpdlpgtnrsgvqarnmldrllelmgqr 234
QY 250 KLDIKEEQEEVEKTEDEAKMSKSDHYRKORTDDLLGWVLKSN----LSTEQILDLI 305
DB 235 k-----esgeft-----dmgylmkknednylltdkeirdqv 267
QY 306 LSLFAGHETSSVALIALAIFLQACPKAVEELREBEETIAAKKEIGSELMWDDYKMD 365
DB 268 vtllysgyeltvstcsmmaikyldhpkaleelrethlairekrp--depltdidksmk 325
QY 366 FTQCVINETLRNGVYRFLRKALKDVRKYGVDIPSGKVLPAVISAHLDSRRYDQNLf 425
DB 326 ftravifetsrlactlvngvlrktchdelngyllpkgyrlyvytrelnydslyedpmif 385
QY 426 NPWRWQOQNNNGASSSGSFSFTWGNINYPFGGPRLCAGSELAKIEMAVFIHLVLRNM 485
DB 386 npwrwmekslesks-----yfllyggvrlcpqkelglsvevssflhytvkyrw 434
QY 486 ELAEDDQPFAPFPVDFPGLPIRVS 510
DB 435 eengeedkimvfprvsapkgyhlkcs 459
```

```
RESULT 13
AAG11835
ID AAG11835 standard; Protein: 462 AA.
XX
AC AAG11835;
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 10707.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
FN EPI033405-A2.
PD
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131048.
PR 30-APR-1999; 99US-0132049.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
```

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-01343768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161921.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 26.6%; Score 712; DB 21; Length 462;
Best Local Similarity 32.5%; Pred. No. 1,1e-54;

Matches 164; Conservative 93; Mismatches 196; Indels 52; Gaps 8;

QY 12 LLLPSSLFLFLILKRRN--KTRPNLPKSGKWPFLGCTGYLKPRTATLGGPMQ 69
DB 2 mmllgllivlvccltalilwqmqmryskgllppgtmwpilgcteflkgp----dlmk 56
QY 70 QHVSQKIVRSNLFGEPTIVSADAGLNRFITONEGRLEFESYRSIGGILGKSMVLVY 129
DB 57 nqrlrysfshlqcpvlvsmdealnrylimeskylvagypqsmldlqtcniaavh 116
QY 130 GDMHRDRSISLNFSLHARLTLILKDVERTLFLDSWQONSIFSQDEAKKFTFNMA 189
DB 117 gprhrlmgslslisptmmkdhllpkldfmrnylqgwdelevdikekkmnaf-1ss 175
QY 190 KHMSMDGEETEQLKKEYTFPMKGVASAPLNPGTAHKAOSRATILKFTERKKEER 249
DB 176 llqaelkpkveeyrteffklvvgtlsvpidipgtnyrsyvgarnidilleteqner 235
QY 250 KLDIKEDQEEVKTDEAMSKSDHYRKQRTDDLLGWLKHSN---LSTEQIIDL 305
DB 236 k-----esgeft-----dmgylmkrednrylltckelrtqv 268
QY 306 LSLFAGHETSSVALAIFLIQACPAVEELREHELEIARAKKELGSELNWDYKMD 365
DB 269 vlllysgvetsstmmaklylhdhpkaleelrhehlairekrp--depltdldiksmk 326
QY 366 FTQCINETLRLGNVRFRLKKAIDVRYKGYDIPSGKVLPIVSAVHLNDSRDQPMLF 425
DB 327 ftravifetsrlativgvlrktchdlnngylipkgwrlvyvcrehlydtslyedpmif 386
QY 426 NPMRMQOONGSASSGSGSFTWGNMNPFGGPRLCAGSLALEMVFTHHVIKFNW 485
DB 387 npwtmekslesks-----ylllgygyrlcpqkelgisevssllhylvlkytw 435
QY 486 ELAEDDQPFAPFPVDFPGLPIRVS 510
DB 436 eengedklmvfprvsapkyhlks 460

RESULT 14

AC AAG11834 standard; Protein; 465 AA.

AC AAG11834.

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 10706.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134378.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139453.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.

PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26.6%; Score 712; DB 21; Length 465;
Best Local Similarity 32.1%; Pred. No. 1,2e-54;
Matches 163; Conservative 94; Mismatches 194; Indels 56; Gaps 8;

OY 14 LIPSLTLFLFLFL-----KRNKRTFRNLPFGSGMPELGTIGYIKPYATATLIGDF 67
DB 3 immilgllivlvcitallrvnqmtyskgyiprglmgwpifgetelfkqgp-----df 57
OY 68 MOOHVSKYKTYRNSLFGEPYVSADAGLNRTLQNEGRLEFCSYPSISGILGKMSMLV 127
DB 58 mkngqrlrygsffkshllgpcflvsmdaelnryllmeskglvagypqsmldlgtcniaa 117
OY 128 LVGDMDRMRSLNPLSHARLRTILLDKDYERHTLVLDMSQOONSIFSAODEKKRTFRL 187
DB 118 vhgpsrhlrmrgslslslsptmmkdhllpklddfrnrylsgwddletvdldqektkhmaf-1 176
OY 188 MAKHMSMDPGEETEQOLKEEYTFPMKGVSAPLNLPGTAYNKALOSRATILKFERRKE 247
DB 177 sslldgaetlkkpeveyteffklvgtlsvpidrglcnysgvagarmldtllelmg 236
OY 248 ERRLDIKEEDQEEVEVKTDEAEMSKSDHVRKQRTDDLLGWLYKHSN---LSTEQILD 303
DB 237 erk-----esgeft-----dmlylmkkehnrnyllttdkeird 269
OY 304 LILSLFAGHESVAIALAIFLOACPKAVEELREHNEIARAKKELOESELMDDYK 363
DB 270 qvvtllsygetvstlsmaalkylhphkaleelrrehalrterkrp--depltdldks 327
OY 364 MDETCVINEETLRLGNVRFELHRKALKADVRKYDIPSGMKVLPVISAHLNDSRPDOPN 423
DB 328 mkfttaivfetsrlativngvlrktthdlnngyllpkvtrlyvrgrelndytslyedpm 387
OY 424 LFNPMRMOONNGASSSGSEFTMGNNYMPFGGPRLCAGSELAKLEMAVFTIHLVLEK 483
DB 388 lfnpmrmmekslesks-----yflfggyvrlpqrkelylsevsflhyvtky 436


```
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 24.4%; Score 653.5; DB 21; Length 468;
Best Local Similarity 30.7%; Pred. No. 1.8e-49;
Matches 160; Conservative 91; Mismatches 189; Indels 81; Gaps 13;

```
QY 9 LPLLLPLSLSLPLLLKRRNRKTRFNLPGKSGMPLGFTIGLKRYTATTLGDFM 68
Db 8 vvpil-----llcillyvyskkkknstgkllp9smgwpylget---lqlysgnp-nviff 59
QY 69 OQHVSKGKTYRNSLFGEPITVSADAGLNRFITONEGRLEFECSPRSIGIILKMSMLVL 128
Db 60 tsbkqkygelfktriljygpvcmlaspeaarfvltvthamtkpkyprskexkligpsalfth 119
```

```
QY 129 VGDHMRDRSISLNFSLHARTLTLKDERHTLFLVDSMOONSIFSADDEAKKTFNL- 187
Db 120 qgdvhsnrlklyqssfyecelrk-lhpdlchialssqswamprlystygemkktfdvg 178
QY 188 ---MAKHMSMDPEEETEDOLKKEYTFMKGVVSAPLNLPGTVMKALOSRAVITLAFIER 244
Db 179 lalifghles----sykeilkhnynlvdkgynsfmslpqtsyhxalmarkqjktlvse 233
QY 245 KMEBRKLDIKEEDQEEVEVTEDEAEKMSKSDHVRKORTDDDLGWLKHSN-----LSTE 299
Db 234 licerr-----ekralqtdfignllntfnekgyvltge 266
QY 300 QIDDLILSLFAGHETSSVALAIFELQACPRAVEELREE---HLEIARAKKELESE 355
Db 267 qiadnllgyvlfagqdtlaasccltwllkylhddqkllveavkeqkalyeansrek-----p 321
QY 356 LNMDDYKKMDFTOCVINETRLGNVVRFLHKKALKDYRYKGYDIPIGSKVLPYISAVHLD 415
Db 322 ltwrqtrmpblthkviveslrmasslftfreaavdvexkylllpkywkmplfrnlhna 381
QY 416 NSRYDQNLFPWRMOQONNGASSSGSFSSTWGNMYPGGGPRLCAGSELAKEAVF 475
Db 382 pkfysnpevfdsrfe-----vnpkpnctimpfsgsvhacpynelaklqllif 428
QY 476 IHHLVLKFNNELAEDDO-----PFAPEYDFPNCGLPIRVS 511
Db 429 lnhlvsnfwevkggekgyldpfpdp---gnqlpatftr 465
```

Search completed: September 28, 2001, 18:44:18
Job time: 48 sec